Mechanism of Transcription and Translation Regulation by Riboswitches in Bacteria

by

Shiba Sundar Dandpat

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
in the University of Michigan
2021

Doctoral Committee:

Professor Nils G. Walter, Chair
Assistant Professor Aaron Frank
Assistant Professor Kristin Koutmou
Associate Professor Jayakrishnan Nandakumar

Shiba Sundar Dandpat

shibas@umich.edu

ORCID iD: 0000-0003-3413-9847

ACKNOWLEDGMENTS

The last five and a half years of my doctoral journey have been the most exciting experience of my career. I am grateful to the University of Michigan, Rackham graduate school, department of chemistry, and the Walter lab for providing me the opportunity to pursue my doctoral journey and helping me every step of the way. I am grateful to so many wonderful people along the way who have supported and influenced me to reach this milestone. First and foremost, I would like to express my gratitude to Prof. Nils G. Walter for his support, encouragement, and mentorship throughout this doctoral journey. I feel lucky to have an advisor like him, who allowed me the freedom to pursue the projects of my interest and provided directions whenever I needed them. I would also like to thank my doctoral committee members (Dr. Jayakrishnan Nandakumar, Dr. Aaron Frank and Dr. Kristin Koutmou) for providing me directions and feedback to complete my dissertation. I am very grateful to my mentors (Dr. Krishna C. Suddala, Dr. Sujay Ray, Dr. Paul Lund, and Dr. Adrien Chauvier) who have provided me direction and coached my skills in and beyond biochemistry and biophysics, which has helped me in putting this dissertation together. I am very grateful to my collaborator Prof. Matt Comstock (MSU), for allowing me to work comfortably and independently in his lab at MSU.

The doctoral journey is undoubtedly a challenging one. I am fortunate to have come across so many beautiful people who helped and supported me through this journey. I feel lucky to be a part of the Walter lab and to have come across some wonderful peers (Surajit, Ameya, Zi, Jieming, Qian, Julia, Nibedita, Poorna, Shankar, Javier, Alex, Kunal, Rajeev, Tanmay, Luhan, Robb, Catie, Karen) who have made my time easier inside and outside the lab and provided a fun and stimulating work environment. I am also grateful to some of my closest friends (Khushboo, Curly, April, Molly, Kwan, Sahil, Saurabh, and Ricardo) with whom I started my doctoral program. They have supported me in every way possible through the ups and downs of the doctoral journey. This journey would have been very difficult without their presence in my life.

This journey would not have been possible without some of my teachers and their faith in me. I would like to thank my teachers and mentors from my undergraduate institute, especially Dr. Moloy

Sarkar and Dr. Himanshu S. Biswal, who believed in my potential to pursue this doctoral journey and supported me in every way possible. Lastly, I would like to express my gratitude to my family, my mother, my sister (Gudia), and my father, whose unwavering support, belief, and sacrifice made my journey possible for me. I hope I have made them proud of this milestone.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Riboswitches are a class of structured non-coding RNA elements that have attracted attention due to their potential as anti-bacterial and anti-microbial targets upon binding to small drug-like metabolites. They are primarily found in the 5’ untranslated region of bacterial messenger RNAs (mRNAs) and regulate gene expression by changing their three-dimensional shape in response to interacting with their cognate ligands (e.g., small molecules and cellular components). This conformational change in the riboswitch’s aptamer influences its expression platform to regulate gene expression by controlling transcription and/or translation regulation. While the mechanism of ligand binding resulting in altering the aptamer structure has been extensively studied, much less is known about how this change influences the downstream expression platform or affects transcription and translation regulation. Using a range of single-molecule and biochemical assays, my doctoral work is focused on understanding riboswitch aptamers and their influence on transcription and translation.

Using these assays, I studied the ligand binding and riboswitch folding mechanism of a Mn^{2+} sensing riboswitch that was known to regulate transcription termination. My work revealed that the riboswitch was sensitive to ligand (Mn^{2+}) binding at sub-millimolar concentration and formed a stable docked conformation required for transcription readthrough only when Mn^{2+} was bound to its ligand-binding core. My observations further demonstrated that mutation of the ligand-binding core and substituting Mn^{2+} ions with other divalent ions did not result in such a stable docked conformation, supporting a specific role of the Mn^{2+} ion for the transcription regulating riboswitch.

To determine the mechanism of riboswitch folding during transcriptional events like pausing of RNA polymerase (RNAP), I next studied the Mn^{2+} sensing riboswitch within a paused elongation complex and observed that both Mn^{2+} and RNAP play essential roles in co-transcriptional folding and the determination of riboswitch conformation. My data further showed

that transcription factor NusA interacting with the transcribing RNA helps in signaling the riboswitch conformation to influence the outcome of transcription.

Next, I studied the ligand-dependent riboswitch folding mechanism of a preQ_1-sensing riboswitch and its role in determining the outcome of translation initiation. Using single-molecule studies, I found that the riboswitch precisely controls the accessibility of the ribosome binding site for 30S binding to the mRNA, a step needed to initiate translation. My data indicated that while preQ_1 binding to the riboswitch both antagonizes 30S recruitment and accelerates 30S dissociation, availability of SD sequence and its distance from the riboswitch aptamer in the mRNA provided hair-triggered precision for 30S binding.

Overall, my work in this dissertation leveraged single-molecule fluorescence microscopy to provide insights into the mechanisms by which riboswitches regulate gene expression at the levels of both transcription and translation beyond their aptamer-ligand interaction.

Chapter 1. Introduction: Riboswitches Observed by Single-Molecule Techniques at the Timescale of Transcription and Translation[1]

1.1 RNA: An essential biomolecule for life

1.1.1 Central Dogma

The central dogma of life provides a pathway for understanding how life exists[1,2]. In the central dogma, DNA is a double-stranded polynucleotides structure and the core genetic blueprint that encodes all genetic information essential for any living organism to survive[1,2]. DNA undergoes a process to transfer fragments of essential information to create a single-stranded polynucleotide called messenger RNA (mRNA) using transcription machinery RNA polymerase (RNAP)[3]. RNA, an intermediate in this process and a carrier of genetic information, further decodes the essential information by using another machinery, called the ribosome, to make proteins that regulate and respond to the downstream processes in the living cell[4]. This flow of genetic information from DNA to make protein through RNA represents an elegant model of how the central dogma defines a pathway for the existence of life.

1.1.2 Role of non-coding RNAs in gene regulation

One of the key players in the central dogma is the intermediate nucleic acid, known as mRNA. mRNA is a member of the RNA family that inherits part of the genetic information of a long DNA genome and decodes to make desired proteins for regulatory pathways[3,4]. However, genome-wide high-throughput sequencing studies later revealed that coding genes making proteins comprise less than two percent of the genetic transcripts derived from a human genome[5,6]. The remaining RNA transcripts were observed to be non-coding genes[7]. Further studies of RNA revealed that a majority, while non-coding in nature, plays key roles in gene regulation. Hence, although non-coding, these RNAs should not be confused with being non-functional[7,8]. The earlier

---

[1] Some contents of this chapter are adapted from a manuscript under the title "Transcriptional riboswitches integrate timescales for bacterial gene expression control" published in *Frontiers in Molecular Biosciences* **2020**, by Catherine E. Scull, Shiba S. Dandpat, Rosa A. Romero and Nils G. Walter.

discoveries of rRNAs and tRNAs and their roles in translation regulation and decoding proteins in conjunction with mRNA and the rest of the translation machinery opened the door for identifying other types of non-coding RNAs and their functional roles[8]. The discovery of RNA enzymes (ribozymes), RNA sensors (riboswitches), miRNAs, RNA introns, lncRNAs, guide RNAs, etc. further demonstrated the diversity of non-coding RNAs and their functional roles at different stages of the gene expression and life-cycle of a cell[8]. In recent years, many types of non-coding RNAs have been identified for their functional role during various diseases like cancer and their application as therapeutics, opening new doors for biomedical applications[6,9,10].

The functionality of non-coding RNA can be attributed to its propensity to fold into a variety of different secondary and tertiary structures on a rugged free-energy landscape[11-14], overcoming internal electrostatic repulsion of its phosphate backbone to form stable hydrogen-bonding between heteroatoms in the form of both Watson-Crick and non-Watson-Crick base pairs and sugar-base interactions[15]. Genome-studies have revealed that non-coding RNAs have a remarkable ability to form robust structures over coding RNA[14,16]. The ability of non-coding RNA to not just adapt stable structures (e.g., RNA hairpins) but also to undergo structural rearrangement in response to external conditions like temperature (RNA thermometers), metabolite concentration (RNA G-quadruplex), and specific ligands (Riboswitches) to mediate gene regulation and achieve critical cellular outcomes, emphasizes their importance in prokaryotes as well as eukaryotes[16,17].

1.1.3 Riboswitches and their role in gene regulation

Riboswitches are a group of non-coding RNA motifs that exist almost exclusively in prokaryotes[18,19]. They are usually found upstream, in the 5' untranslated region, of mRNAs where they control key gene regulatory processes like transcription and translation[19]. They consist of two main parts: an aptamer domain involved in ligand sensing and an expression platform domain involved in modulating structural change in response to ligand binding[19]. Riboswitch aptamers have evolved to bind diverse ligands, which often cellular metabolites, ranging from small molecules (often metabolites like s-adenosyl-methionine and preQ$_1$), to cations (such as Mn^{2+}), anions (such as F$^-$), and even other RNAs (such as tRNAs)[20-24]. Upon binding of a cognate ligand to the aptamer, riboswitches alter the secondary structure of a downstream expression platform, which either: 1) results in alterations in transcription termination, or 2) disallows translation initiation (Figure 1.1)[25]. Till date, there are more than 41 validated classes of riboswitches based on the types of ligand they recognize, spanning more than 100,000 riboswitch representatives[26]. Even after such large diversity in riboswitch, the list is still underrepresented due to the existence of orphan riboswitch candidates whose cognate ligand has not been identified yet[27,28]. This growing list of riboswitch candidates to determine their mechanism for identifying and interacting with ligand and how this mechanism correlates with the control of gene regulation makes riboswitch exciting regulatory machinery to study.

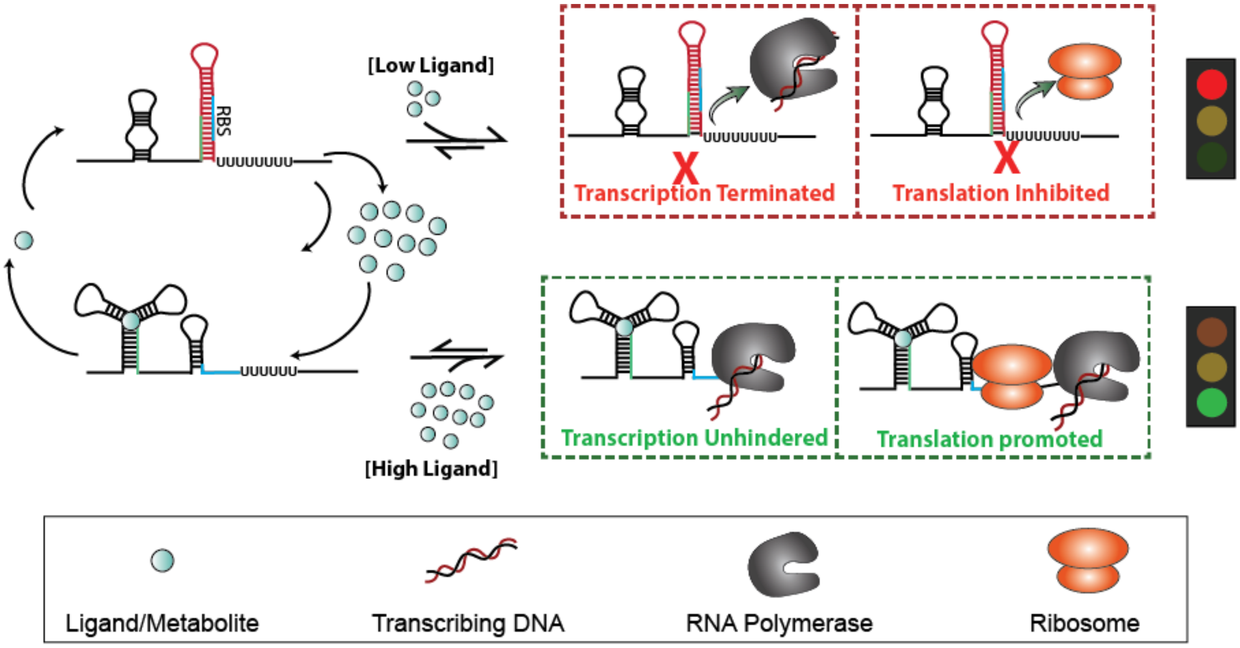


Figure 1.1 General mechanism of action for riboswitch in regulating transcription and translation.

Riboswitches work in a feedback loop (left) in which they rearrange their structural orientation of the aptamer depending on the concentration of ligand. This ligand-dependent structural rearrangement in aptamer affects the expression platform (right) by either downregulating or upregulating gene expression reflected on the transcription and translation processes. Riboswitches controlling transcription switch between transcription termination or elongation dependent on ligand concentration to control gene expression. Riboswitches controlling translation regulation switch between translation initiation inhibition or promotion dependent on the ligand concentration.

1.1.4 Riboswitch-mediated genetic control mechanism

Riboswitches control gene regulation by precisely sensing a cognate ligand, which is often a cellular metabolite. What makes a riboswitch so unique is its ability to precisely discriminate different cognate metabolites at high precision that ultimately determines the fate of the downstream gene. It does so by working in a regulatory feedback loop depending on the concentration of ligand required by the cell under specific conditions[29]. When the concentration of an essential metabolite is low, and the bacterium needs to increase the availability of that metabolite, the riboswitch adapts a structure that can turn ON a gene expression pathway to let this process happen. When the metabolite concentration gets too high, the riboswitch allows them to bind its core and change the structural rearrangement to turn ON certain genes, enabling the release of excel metabolites out of the cell to attain the condition as before. This precisely tuned feedback loop regulating gene expression in response to ligand concentration makes a riboswitch an impressive regulatory system in bacteria and helps them survive harsh environmental conditions with limited cellular resources. The two primary ways in which riboswitches control gene expression are through transcription and translation.

Upon binding of a ligand to the aptamer during transcription regulation, the riboswitch controls the expression platform by either promoting or disrupting transcription termination[19,25]. While promoting transcription termination, the riboswitch adopts an alternate aptamer structure that results in forming a stable terminator hairpin in the expression platform[19,25]. In contrast, while enabling transcription readthrough, it switches between aptamer conformations that disrupt the terminator hairpin or forms a competing secondary structure, known as anti-terminator[19,25]. Formation of a terminator structure in response to ligand binding to the riboswitch signals downregulation of associated genes that trigger RNAP to dissociate from the transcribing DNA template[25]. Disruption of the terminator structure or formation of an anti-terminator in response to ligand binding to the riboswitch signals upregulation of gene that signals RNAP to continue transcribing[25]. Some transcription-regulating riboswitches use Rho-dependent terminators instead of intrinsic terminators to signal the RNAP to stop[25].

During translation regulation, the riboswitch in response to ligand binding decides the fate of ribosome binding to initiate translation by controlling the accessibility of the ribosome binding site[19,25]. During downregulation of translation, the riboswitch adopts a structure that sequesters the ribosome binding site, also known as the SD sequence in the mRNA of bacteria. Such sequestration of the SD region prevents the 30S ribosome from binding to an mRNA and initiating translation[19]. By contrast, to upregulate translation the riboswitch alters its structure to unsequester the SD region, allowing the 30S ribosomal subunit to recognize the ribosome binding site, essential for initiating SD mediated translation[19,25].

1.2 Timescale with the mechanism of action of riboswitch

1.2.1 Timescale of riboswitch folding and its correlation with gene regulatory outcomes

The mechanism of action of riboswitch can essentially be broken down into 4 steps that occur on distinct timescales: 1) ligand influx, which can be as fast as the rate of diffusion, 2) ligand binding to the RNA, establishing specific hydrogen bonding, stacking and ionic interactions (nanoseconds to single seconds), 3) alterations in RNA secondary structure (milliseconds to multiple seconds), and 4) regulation of transcription elongation/termination or translation initiation/inhibition and their downstream biological consequences (seconds to minutes)[30].

Studies on structural and conformational dynamics of riboswitches have been extensively performed on the riboswitch aptamer using a wide range of biophysical methods[31,32]. In reality, that riboswitch folding *in vivo* occurs co-transcriptionally that ultimately decides the fate of expression of downstream genes[33], leaving a gap between the molecular and cellular levels of understanding. RNA structures have recently been revealed to impact active site conformations and transcription activity of bacterial RNAP in both cis and trans[34,35]. Recent work on $\mathrm{preQ_1}$-sensing riboswitch unveiled that the nascent riboswitch's secondary structure directly influences the pausing behavior of the transcribing RNAP, illustrating that both the template DNA and RNAP have a significant impact on riboswitch folding, and vice versa[23]. Similarly, to attain precise control of the translational pathway in terms of ribosome binding, intricate folding of the riboswitch aptamer and its influence on expression platform containing RBS must be achieved on a timescale relevant to translation events[36]. Thus, to gain proper control of transcript synthesis and ultimately protein expression, the four steps of riboswitch activity, which exist in distinct timescales, must be kinetically coupled[37].

1.2.2 Correlating the timescales of ligand binding to riboswitch folding

During riboswitch control of gene expression, the binding rate of the ligand and the folding rate of the RNA may be faster (nanosecond to millisecond timescales) than the rate of transcription (milliseconds to seconds)[19,38-41]. Hence the cellular concentration of ligand must be higher than its dissociation constant (K_D) to allow for the ligand-binding kinetics to outpace and thus drive the kinetics of RNA folding. A commonly observed mechanism called the induced fit (IF) mechanism describes riboswitch ligand-binding events where ligand binding occurs faster than the conformational changes of the riboswitch, whereas ligand dissociation from the unfolded RNA is slower[40]. Alternatively, the conformation selection (CS) model describes riboswitches where the RNA conformational change occurs faster than ligand binding, and ligand dissociates too rapidly from the unfolded RNA to achieve IF mechanism[22]. The transition between the IF and CS models is governed, on the one hand, by the ligand concentration and, on the other hand, by temperature and cofactors affecting RNA folding such as the cationic micro-environment. Undoubtedly, evolutionary pressures shape the sequence composition of the riboswitch to finetune this balance to the cell's needs[22,42-44]. Ligand recognition mechanisms like the CS and IF models have provided the bases for the kinetic selection of transcriptional riboswitches[43,44].

Modern biophysical techniques such as single-molecule fluorescence resonance energy transfer (smFRET) and optical tweezers have quickly become essential tools for monitoring the structural changes in the aptamer region in response to cognate ligand binding[32]. smFRET in particular, has been widely used to probe conformational changes in riboswitches at varying concentrations of ligand to correlate ligand binding with RNA folding[32,37,44]. For example, Holmstrom *et al.* investigated the unique example of the hydroxocobalamin (HyCbl) riboswitch, where binding of HyCbl can be monitored by ligand-induced quenching of fluorescence from the donor and acceptor fluorophores and unambiguously linked to RNA conformational changes[45]. In other cases, such as *in-vitro* evolved aptamers with more open binding pockets, the ligand itself can be labeled to monitor its binding to single RNA molecules[46]. In most reports, however, indirect changes in RNA folding and unfolding dynamics are used to probe the mechanism of the aptamer-ligand interaction. For example, one of the smallest riboswitches, the class I preQ$_1$-sensing riboswitch, follows the two ligand-binding mechanisms of IF and CS independent of the ligand and metabolite conditions as well as specific sequence adaptations found in various bacteria[22,43]. The transcriptional preQ$_1$ riboswitch from *Bacillus subtillis* (*Bsu*) has been observed to favor a CS pathway where the ligand primarily binds to a pre-selected conformation of the aptamer[43]. However, the ligand recognition mechanism is fluid, as the same riboswitch can adopt instead the IF mechanism at low metabolite and Mg^{2+} concentrations[22]. This dependence of the folding pathway on the relative timescales of ligand binding and conformational dynamics of the aptamer can be identified as a kinetic coupling mechanism occurring early in the decision tree of gene regulation. Similar kinetic control mechanisms of ligand recognition by the aptamer have been observed to be operational in multiple riboswitches[42,47-49].

1.2.3 Correlating the timescales of riboswitch folding to the transcription control of gene expression

During bacterial transcription, both the kinetics of ligand binding and the speed of RNA transcription determine the functionality of the riboswitch beyond the ligand binding to its aptamer[39]. The coupling observed for the rates of transcription elongation and RNA folding as it emerges from the RNAP exit channel in 5’-to-3’ direction highlights the importance of studying riboswitches in the context of the transcription elongation complex. In fact, it has been demonstrated in other RNA folding systems (such as bacterial ribosome biogenesis) that the cotranscriptional directionality of its folding influences an RNA’s interactions with known binding partners[50,51]. Transcriptional regulatory events, such as pausing, are crucial to the balance between RNA folding and the speed of additional RNA sequences emerging in the wake of RNAP[52]. Studies have shown how sequence-specific pausing allows the nascent RNA to reach an equilibrium of folded states that then can be further stabilized by RNA binding molecules[23,38,51]. Depending on cellular conditions, transcription factors such as NusA and NusG are found to stabilize and disrupt transcriptional pausing, respectively[53-55]. During transcription elongation, a cascade of faster events, including RNA folding, ligand binding, and interactions of the RNA with RNAP, together with variations of transcription speed over time, integrate over the biological state of the cell to govern downstream gene regulation. In light of the reversibility of many, and irreversibility of some, of these steps, conformational and kinetic proofreading becomes possible, adding critical layers of control over the ultimate gene expression outcome[56].

Several recent studies have highlighted details of the cotranscriptional nature of riboswitch folding. High-resolution optical tweezers and single-molecule force spectroscopy approaches showed that a co-transcriptionally folded adenine riboswitch undergoes transcription readthrough predominantly in the presence of adenine, while its absence leads to transcription termination[33]. smFRET assays further demonstrated kinetic control of cotranscriptional folding of a thiamine pyrophosphate (TPP) riboswitch[57]. The isolated riboswitch aptamer was observed to fold into a translation “off” conformation independently of its TPP ligand. By contrast, transcriptional pausing allows the riboswitch to rearrange into an “on” conformation in the absence of TPP, while ligand binding steers the nascent RNA into the “off” conformation to downregulate gene expression. This work illustrated that only a brief time window between transcriptional pausing and ligand binding determines the fate of downstream gene expression[57]. A combination of smFRET, biochemical studies and, and molecular dynamics simulations on $preQ_1$ riboswitch have shown that on the time scale of transcription, pausing allows time for the RNAP to slow down and the riboswitch to sense ligand, ultimately stabilizing a fully folded RNA pseudoknot conformation and releasing the paused RNAP[23].

1.2.4 Correlating the timescales of riboswitch folding to the translation control of gene expression

During bacterial translation regulation, occlusion or liberation of the RBS (also known as SD sequence) in response to riboswitch folding decides the fate at the initial stages of ribosome binding to an mRNA, hence ultimately determining the fate of protein expression[36]. The transcribed mRNA competes to initiate translation for expression of protein over the competing degradation process, making the necessary folding of RNA tightly regulated and kinetically driven for efficient translation[37,58]. For precise control of translation initiation by a riboswitch, an intricate conformational rearrangement of the riboswitch in response to ligand recognition must occur within the timescale of events during the initial stages of translation. Besides the riboswitch, multiple additional factors during these initial stages of translation initiation, such as the SD sequence, ribosomal subunits and translation factors, influence the initiation rate and tune the extent of protein expression for a riboswitch regulated gene[36].

The SD sequence is a highly conserved purine-rich sequence of 6-8 nucleotides located upstream of the start codon[59]. During SD-mediated translation, the 30S ribosomal subunit scans for and, upon recognition, pairs with the SD sequence using a complementary and conserved anti-SD sequence in its 16S rRNA component, initiating ribosome recruitment[60]. The degree of complementarity of the SD:anti-SD base pairing interaction has been found to be a key factor for determining the recruitment and loading of 30S subunits to mRNA[61,62]. Mutation and reporter gene assays have shown that disruption of the SD:anti-SD interaction drastically influences the level of protein expression[61,63,64]. Using a single-molecule colocalization-based SiM-KARTS (Single-Molecule Kinetic Analysis of RNA Transient Structures) assay, Rinaldi *et al.* showed that a translation regulating $preQ_1$ riboswitch can influence the accessibility of the SD region to binding a complementary probe carrying the anti-SD sequence[65]. However, the roles of the SD:anti-SD interaction in the native context of a riboswitch binding the ribosome and in the resulting translational outcome are still unclear.

Ribosomal protein S1 is an RNA binding protein that is known to unwind double-stranded RNA[66]. It is also a key translational factor that allows the binding of 30S subunit to mRNA through its RNA chaperone activity, helping to unfold of structured mRNAs and assisting the 30S for its correct positioning with the initiation codon for translation[67]. NMR and smFRET studies have provided mechanistic insights into the role of S1 protein and revealed that S1 binds to the structured part of the translation initiation region and melts them to make them possibly accessible to the 30S recognition[68,69]. However, little is known about how the role of S1 facilitates 30S recognition and enables translation initiation of structured mRNAs.

1.3 Single-molecule methods to study the mechanism of action of riboswitches

1.3.1 Single-molecule Förster Resonance Energy Transfer (smFRET)

Single-molecule FRET (smFRET) is a well-known biophysical technique that enables precise measurement of the distance between a pair of molecules (often fluorescent organic dyes, but they can be quenchers, nanoparticles or quantum dots) termed a Donor-Acceptor pair. It does so by non-radiative transfer of energy from the donor molecule, when excited, to the acceptor molecule and results in emission from the acceptor[70,71]. The efficiency of such energy transfer (E), which can be measured, is dependent on the distance (r) between the Donor-Acceptor pair, hence making smFRET a powerful tool and often referred to as "molecular ruler"[72]. This efficiency of energy transfer, E, is given as $E = (1 + (R / R_0)^6)^{-1}$, where R is the inter-dye distance, and R_0 is the Förster distance when transfer efficiency is 50% (i.e., E = 0.5). smFRET has been extremely powerful to measure distance with the range of 1-10 nm between the dye molecules[70,71]. The energy transfer efficiency shows that when the donor and acceptor are close, the transfer efficiency (E) is maximum and when they are far, the transfer efficiency is minimum.

Over the past few decades, smFRET has emerged as a powerful tool to measure conformational dynamics in biomolecules such as DNA[73,74], RNA[22,43,44], and proteins[75,76] when strategically labeling specific positions. It has recently emerged as a powerful tool to measure interactions between different biomolecules as well as their conformational changes[70,77,78] (Figure 1.2A). Advancements in smFRET for measurements beyond simple bimolecular (two-color) energy transfer[70], like three and four-color FRET and parallel two-color FRET, have enabled measurement of the dynamics of complex biological systems like transcription[57], translation[79], and splicing[80] at the single-molecule level and high precision.

While smFRET is extremely powerful in measuring the conformational dynamics and interaction within biomolecules, it has some limitations. A major limitation is the need for site-specific fluorophore labeling within the FRET distance range, which can be expensive and of low efficiency for large biomolecules like RNAs and proteins when aiming to add fluorophores at the right place[70]. A second limitation of smFRET is the risk of photobleaching, limiting the length of the observation window. Either of the fluorophores being prone to photobleaching limits the time over which smFRET can be measured. While smFRET with wide-field microscopy allows monitoring a large number of molecules at the same time at the expense of photobleaching, the use of confocal microscopy allows monitoring one molecule at a time, allowing high signal-to-noise and preventing photobleaching elsewhere in the sample. The choice of fluorophore pair, deployment of an oxygen-scavenging solution, and budgeting of photons to limit photobleaching (by lowering the excitation intensity or duration) are strategies for studying biomolecules by smFRET for an extended period of time[70].

1.3.2 Single-Molecule Kinetic Analysis of RNA Transient structure (SiM-KARTS)

SiM-KARTS is a fluorescence colocalization based technique that addresses some of the limitations of smFRET and enables measuring conformational dynamics of biomolecules and their intermolecular interaction[20,65]. Inspired by DNA-PAINT[81] based super-resolution technology, it has been implemented to study RNA structure and interaction dynamics (Figure 1.2 B). While it requires fluorophore labeling, unlike smFRET, they are external to the molecule of interest. In addition, since it relies on the kinetics of binding and dissociation of a fluorophore-labeled probe, it is not heavily influenced by photobleaching, allowing an observation window from a few to tens of minutes. Using SiM-KARTS, Rinaldi *et al.* demonstrated a role for a preQ$_1$ sensing riboswitch embedded in the 5'-UTR of a native mRNA in regulating the accessibility of the SD sequence[65]. In addition, Suddala *et al.* have used SiM-KARTS assay to demonstrate the role of the stem (P1.1) and its stabilization in an Mn^{2+} sensing riboswitch for determining the fate of transcription[82]. The ability of the SiM-KARTS assay to detect changes in riboswitch secondary structure without the need for site-specific fluorophore labeling of the target and with an in principle infinite observation window makes it a powerful tool to monitor structural change in virtually any RNA under a wide range of conditions, poising it for a broad range of applications.

1.3.3 Prism based Total Internal Reflection Fluorescence (P-TIRF) microscope for single-molecule imaging

In a TIRF microscope, an evanescent field is created by the reflection of the excitation light source and extends up to ~100-200 nm from the surface to generate wide-field illumination (Figure 1.2C). Due to such limited surface excitation, any fluorophores in the solution background are extremely rarely excited, greatly increasing the signal-to-noise ratio for fluorescence measurements, as is essential for single-molecule imaging[70]. The evanescent field is generated using laser sources for up to four excitation colors at 473 nm (Blue), 532 nm (Green), 639 nm (Red), and 730 nm (IR)[83]. A quartz prism is used to focus the laser beam with an incidence at the critical angle for TIR such that the sample placed at the surface is exposed to the evanescent field at the quartz-water interface for excitation. Upon surface excitation, emission intensities of hundreds of molecules across the field of view can be recorded simultaneously for tens of minutes of observation window. Using sCMOS camera technology, events can be recorded at a time resolution of milliseconds to seconds.

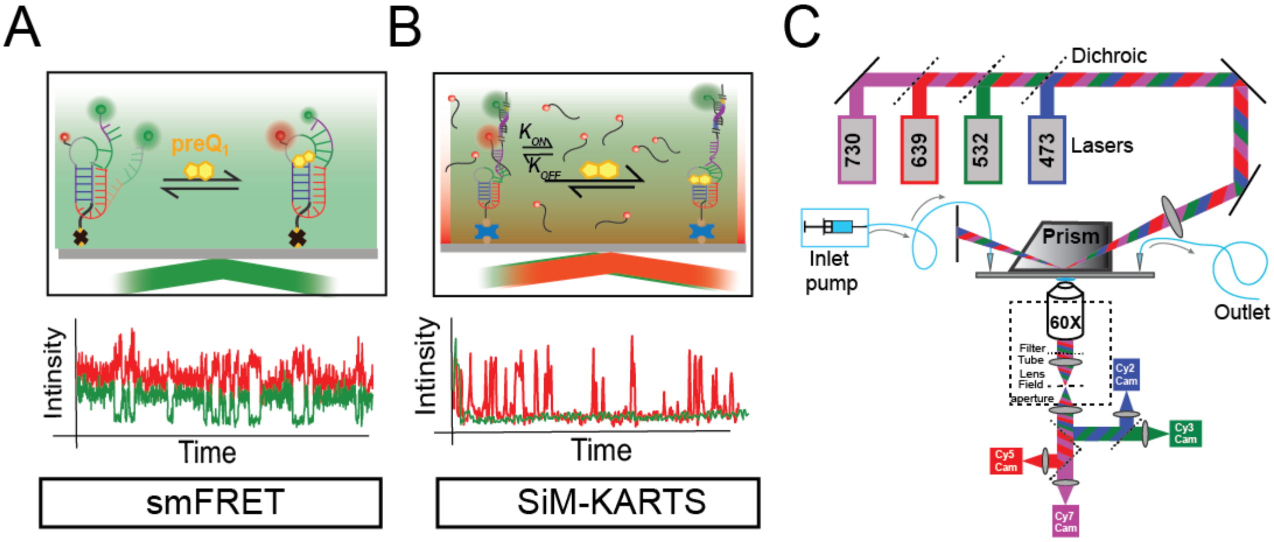


Figure 1.2 Illustration of single-molecule assays and Prism-TIRF microscope.

(A) smFRET assay (top) monitors conformational dynamics of a riboswitch. The intensity (bottom) of donor (green) and acceptor (red) signals are anti-correlated and used to calculate transfer efficiency or FRET efficiency ($E = I_A / (I_D + I_A)$), which gives the distance between donor and acceptor molecules and conformational change associated with the distance. (B) SiM-KARTS assay (top) monitors RNA structure and interaction dynamics. The Cy5 intensity used to monitor interaction dynamics through the binding of a probe based on RNA structure. The Cy3 intensity used for the colocalization of two signals and ensure specific binding. (C) A four-color excitation-based Prism-TIRF (p-TIRF) microscope was used for the single-molecule assays. For most assays probing riboswitch conformational dynamics, only two-color excitation using 532 nm and 639 nm laser excitation sources were used. Although a combination of two or more excitations can be used for a combination of SiM-KARTS and smFRET based assays to monitor complex long-range interactions or RNA-protein interactions.

The data collected from the imaging experiment are post-processed using well-known single-molecule analysis software like HaMMy[84], VBFRET[85], EBFRET[86], SPARTAN[83] or custom-built MATLAB codes to extract and model single-molecule data to obtain kinetics and dynamic information. The prism-based TIRF microscope has been widely used for studying the conformational dynamics of a variety of non-coding RNAs, such as ribozymes[87,88], riboswitches[22,23,43,44,82], RNA G quadruplexes[89], long non-coding RNAe[90] etc. It has also been used to study complex RNA-protein based systems like transcription[23,91], translation[92,93], the spliceosome[94,95] etc.

1.4 Dissertation Goal

This dissertation consists of five chapters. The first chapter provides an introduction about non-coding RNA elements and their role in gene regulation, with a specific emphasis on riboswitches. It highlights the role of riboswitch-mediated gene regulatory processes in bacteria and explains how they control transcription regulation and translation initiation within the timescale of gene regulation. It then provides an overview of the role of timescale at which ligand recognition to riboswitches occurs and further correlates it with regulatory processes, particularly transcription (transcription pausing and cotranscriptional RNA folding) and translation (SD regulated ribosome recognition to mRNA and translation initiation) in bacteria.

Chapter 2 focuses on a transcription regulating riboswitch and is adapted from a manuscript published in *Nature Communications*[82], entitled “Local-to-global signal transduction at the core of a Mn^{2+} sensing riboswitch” for which I am the second author. This chapter identifies the ligand-binding mechanism of a Mn^{2+} ion sensing riboswitch and correlates this mechanism with transcription control using single-molecule techniques (smFRET and SiM-KARTS), and the observations were corroborated with x-ray crystallography and MD simulations.

Chapter 3 presents a mechanistic study of the ligand recognition and conformational dynamics of a Mn^{2+} sensing riboswitch observed at and beyond a transcriptional pause, aiming to understand the correlation between transcriptional pausing and riboswitch folding. Using single-molecule techniques (smFRET and SiM-KARTS) in equilibrium folding and cotranscriptional folding assays, corroborated by biochemical assays, it explores the role of RNAP and transcriptional factor NusA (known to stabilize pausing) alongside ligand Mn^{2+} on determining the role of the riboswitch in controlling transcription.

Chapter 4 focuses on a translation-regulating riboswitch and is adapted from a manuscript to be submitted soon, entitled "Precise tuning of bacterial translation initiation by non-equilibrium 5'-UTR unfolding observed in single mRNAs", for which I am the second author. This chapter identifies the role of a $preQ_1$-sensing riboswitch in regulating ribosome binding and recognition required for translation initiation. It dissects the roles of SD sequence, ribosomal protein S1 and initiation factors alongside the riboswitch aptamer in determining the fate of translation.

Chapter 5, the final chapter recapitulates the general conclusions from each chapter with an extended discussion of future directions for this dissertation.

Chapter 2. Local-to-Global Signal Transduction at the Core of a Mn^{2+} Sensing Riboswitch[2]

2.1 Overview

The widespread Mn^{2+}-sensing *yybP-ykoY* riboswitch controls the expression of bacterial Mn^{2+} homeostasis genes. However, the mechanism of this exquisite metal ion sensing riboswitch is poorly understood. Here, we first determine the crystal structure of the ligand-bound *yybP-ykoY* riboswitch aptamer from *Xanthomonas oryzae* at 2.96 Å resolution, revealing two conformations with docked four-way junction (4WJ) and incompletely coordinated metal ions. Using single-molecule FRET, we unveil a previously unobserved extended 4WJ conformation that samples transient docked states in the presence of Mg^{2+}. Only upon adding sub-millimolar Mn^{2+}, however, can the 4WJ dock stably, a feature lost upon mutation of an adenosine contacting Mn^{2+} in the core. These observations illuminate how subtly differing ligand preferences of competing metal ions become amplified by the coupling of local with global RNA dynamics.

2.2 Introduction

Riboswitches are structured RNA domains commonly found in the 5'-untranslated regions of bacterial mRNAs, where they regulate many essential and virulence genes in response to binding of a specific ligand[26,96,97]. Currently, there are over 40 different riboswitch classes known to respond to ligands ranging from metabolites[98], enzyme cofactors[99], signaling molecules[100-102], tRNAs[103], and to metal ions[104-106]. Ligand binding generally stabilizes a conformation of the riboswitch that modulates either Rho-independent transcriptional termination or translation initiation through accessibility of the Shine-Dalgarno (SD) sequence. The static ligand-bound

---

[2] The contents of this chapter are adapted from a manuscript of the same title published at *Nature Communications* **2020**, by Krishna C. Suddala, Ian R. Price, Shiba S. Dandpat, Michal Janeček, Petra Kührová, Jiří Šponer, Pavel Banáš, Ailong Ke, and Nils G. Walter. K.C.S and I.R.P. are co-first authors. K.C.S., I.R.P., S.S.D., A.K. and N.G.W. designed the experiments; P.B. and J.Š. designed the MD simulations; K.C.S. and S.S.D. performed the single-molecule experiments and analyzed the data; I.R.P. and A.K. performed the X-ray crystallography experiments and analyzed the data; M.J., P.K. and P.B. performed and analyzed the MD simulations; K.C.S., I.R.P., A.K., P.B. and N.G.W. wrote the bulk of the paper. All authors read and commented on the paper.

structures, and the ligand-recognition modes, of a number of riboswitch aptamers have been determined at atomic resolution[44,97,106,107]. Often, the ligand occupies a linchpin position in the global fold where distal residues of the RNA are brought together; however, the dynamic paths by which the local binding of a ligand as small as a metal ion are transduced into the large-scale molecular rearrangements necessary for a regulatory decision by the gene expression machinery largely remain enigmatic[107]. The *yybP-ykoY* RNA motif is one of the most widespread riboswitches across bacteria, including many human and plant pathogens[19,108,109]. It has evolved to sensitively detect Mn^{2+} metal ions and broadly regulate a variety of genes, particularly those involved in Mn^{2+} homeostasis, at the levels of either transcription or translation[106,108,110]. We and others have previously solved crystal structures of the aptamer domain of this riboswitch[106,111], henceforth simply referred to as "riboswitch", revealing that it senses the charge, geometry, and Lewis-acid hardness of Mn^{2+} by forming direct inner-sphere contacts from five phosphoryl oxygens and the N7 of an invariable adenosine. The global structure showed that formation of the Mn^{2+} binding site requires "docking" of two distal helical legs of a four-way junction (4WJ) to form a paperclip-shaped global architecture, facilitated by an A-minor interaction and a second, nonspecific divalent metal binding site (Figure 2.1)[106]. However, the transduction of ligand binding in this metal-sensing core into global structural changes that affect the distal helix P1.1 involved in riboswitching is not understood, rendering it an archetypical representative of our level of understanding of many crystallized riboswitches[107].

The crystal structures of two ligand-bound states (Appendix Figure 6.1 & Table 6.1) of the *yybP-ykoY* transcriptional riboswitch from the rice pathogen *X. oryzae*[112,113], captured in distinct conformations offer snapshots of structural changes *en route* to full ligand binding. Using single-molecule FRET (smFRET), we investigate the global structural dynamics of the riboswitch in the presence of varying concentrations of Mg^{2+} and Mn^{2+}, as well as other transition metals, revealing a previously unobserved extended conformation of the riboswitch. We show that addition of Mg^{2+} induces two kinetically distinct docked and undocked conformations that are in dynamic equilibrium with one another. In contrast, upon addition of sub-millimolar Mn^{2+}, the riboswitch adopts a stable docked conformation that becomes abolished upon mutation of the conserved core adenosine. Taken together, our work reveals the ligand-dependent (un)folding pathway of the Mn^{2+} sensing riboswitch as a guide for how subtle binding preferences distinguishing two similar metal-ion ligands cascade through the coupling of local with global RNA conformational dynamics into powerful effects on gene regulation.

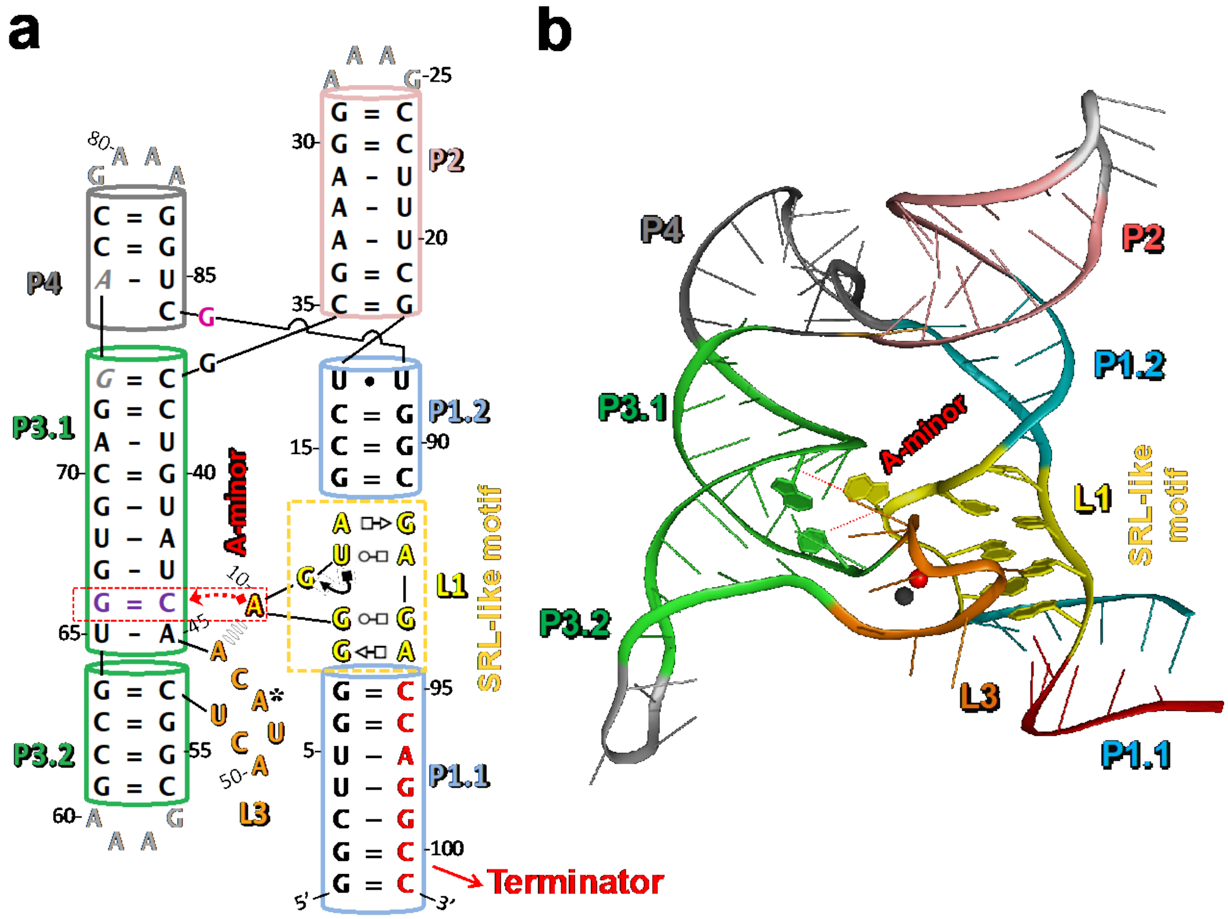


Figure 2.1 Sequence and structure of the *X. oryzae* (*Xory)* crystal structure.

(**a**) Secondary structure of the *X. oryzae* crystal construct. A native CA dinucleotide was omitted between G73-A74 (gray italics) for crystallization purposes only. The A-minor interaction between L1 and L3 and the SRL-like conformation of L1 are shown. (**b**) Crystal structure of a riboswitch conformer showing different secondary structures.

2.3 Results

2.3.1 smFRET reveals an undocked state that docks upon Mg^{2+} addition

To probe the global structural dynamics in the presence of Mg^{2+} and Mn^{2+}, we used smFRET assay to monitor fluorophores positioned on the distal legs of the *Xory* riboswitch (Figure 2.2, Methods). The two RNA strands of the riboswitch with two fluorophores in each are annealed together and attached to the quartz slide surface though streptavidin-biotin interaction for single-molecule imaging (Methods).

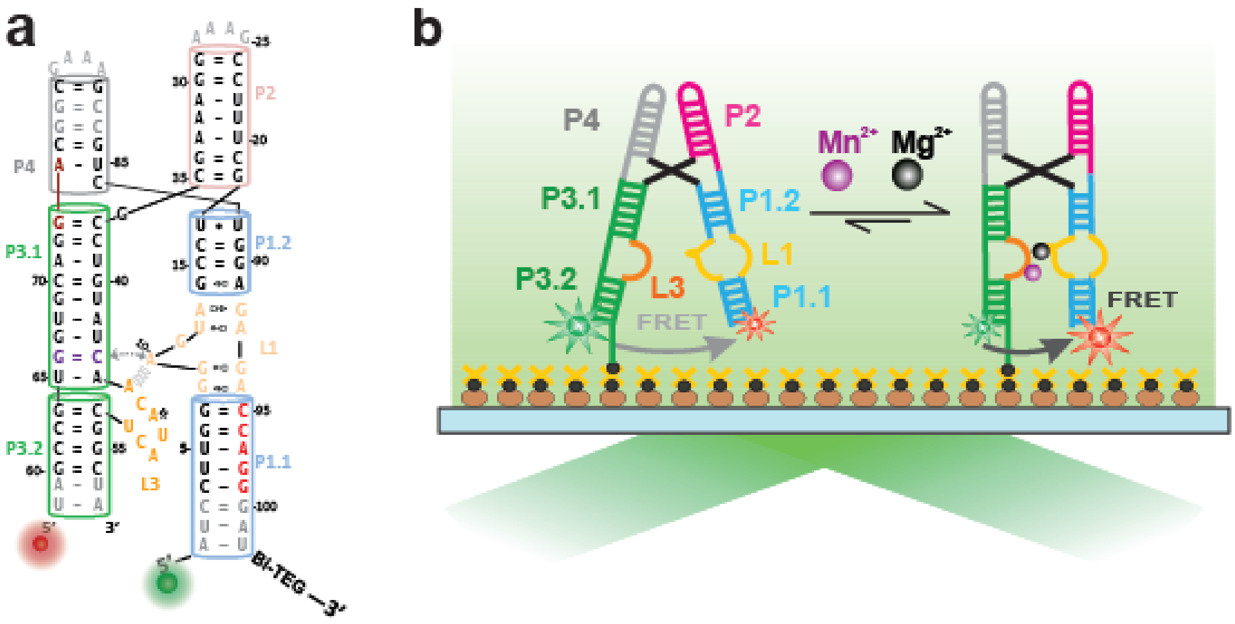


Figure 2.2 smFRET construct and assay for Mn^{2+} sensing riboswitch.

(a) Sequence and secondary structure of the *Xory yybP-ykoY* riboswitch construct used for smFRET. The construct was made was hybridizing two chemically synthesized RNA oligonucleotides modifications at their 5' and 3' ends with Cy3, Cy5 fluorophores and Biotin-TEG. (b) Schematic of the smFRET experiment using p-TIRF microscope indicating the fluorophore labeling positions on the riboswitch.

smFRET traces at 100 mM KCl in the absence of any divalent metal ions showed a stable low-FRET value of ~0.1 without global dynamics (Figure 2.3a), with a population FRET histogram displaying a single peak centered on ~0.13±0.10 (mean ± standard deviation) (Figure 2.3b). The non-dynamic nature of the traces is also evident as an 'on-diagonal' contour centered at ~0.13 in the transition occupancy density plot (TODP), which represents as a heat map the fraction of single molecule traces that exhibit any given specific initial-to-final FRET transition at least once (Figure 2.3c and Methods)[114]. This FRET value corresponds to an estimated distance of ~74 Å between the two fluorophores and suggests an extended, stably undocked (SU) conformation where the two RNA legs are distal and do not interact, unlike the docked crystal structure (Figure 2.1).

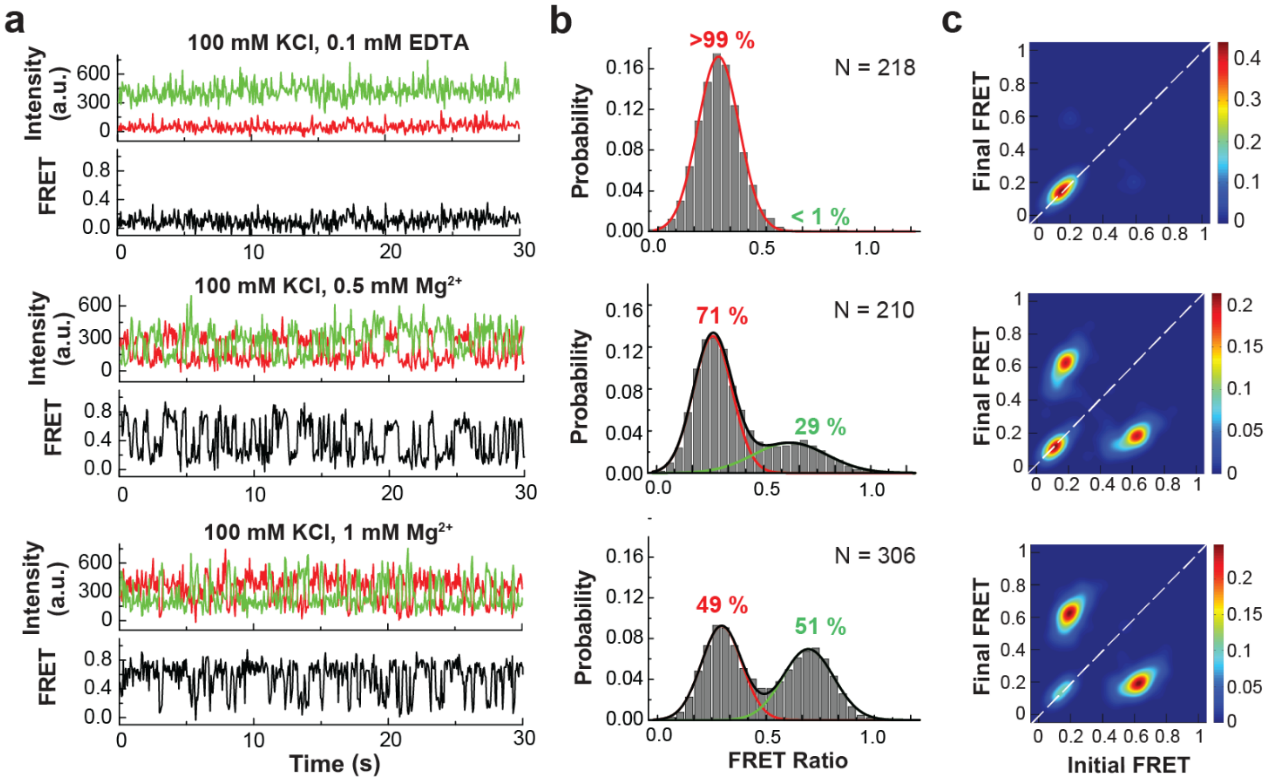


Figure 2.3 smFRET analysis of the WT *Xory* Mn^{2+} riboswitch.

(**a**) Representative smFRET traces under different buffer conditions (top to bottom): absence of divalents (+0.1 mM EDTA), 0.5 mM Mg^{2+} and 1 mM Mg^{2+}, respectively. Green, Cy3; Red, Cy5; Black, FRET. (**b**) Population FRET histograms showing the equilibrium distribution of two FRET states under the conditions of panel (**a**). Gaussian peaks for the low- and high-FRET states are shown in red and green, respectively with the cumulative fit shown in black. Reported are the percentages of FRET states at equilibrium, as well as the number of molecules N analyzed. (**c**) TODPs showing the 'static' and 'dynamic' trances as 'on-diagonal' and 'off-diagonal' heat map contours, respectively. The color code indicates the fraction of each population.

Addition of Mg^{2+} up to 0.1 mM did not result in significant changes in the FRET histograms since almost all traces remained in the SU conformation, with ~3 % of them showing brief excursions into a higher ~0.6-FRET state (Figure 2.4a). Further raising the Mg^{2+} concentration resulted in more dynamic traces transiently adopting this high-FRET state, accompanied by a corresponding decrease in the population of SU traces (Figure 2.4b).

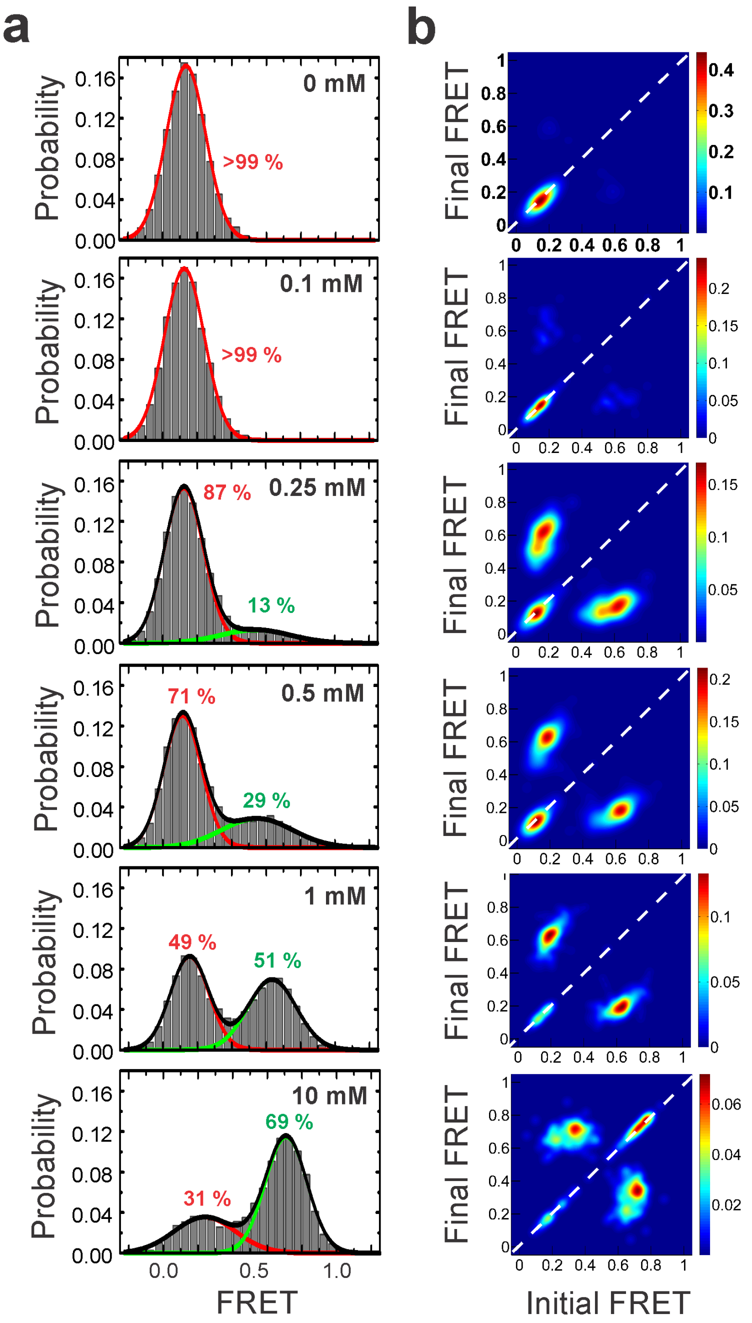


Figure 2.4 Titration of Mg^{2+} at different concentration with the WT *Xory* riboswitch.

(a) FRET histograms showing the distribution of the two FRET states under various Mg^{2+} concentrations, fit to a sum of Gaussian functions. The Gaussian peaks for the low- and high-FRET states are shown in red and green, respectively, while the cumulative fit is shown in black. The histograms under 0 mM and 0.1 mM contain very low populations of the high-FRET state and were fit to a single Gaussian function. (**b**) TODPs under different Mg^{2+} concentrations showing the fraction of static 'on-diagonal' and dynamic 'off-diagonal' molecules. The SD populations under the high, 10 mM Mg^{2+} is evident in the TODP.

At a near-physiological concentration of 1 mM Mg^{2+}, the time- and population-averaged distribution between low- and high-FRET, with mean FRET values of ~0.15±0.11 (49 %) and 0.63±0.14 (51 %), respectively, became almost equal (Figure 2.3b & 2.4a). A FRET value of 0.63 corresponds to a distance of ~49 Å between the two-labeled RNA arms, similar to the distance observed in the crystal structures, suggesting adoption of the compact 'docked' conformation. Reaching 10 mM Mg^{2+}, the fraction of this docked conformation further increased and saturated at ~69 %, with a sigmoidal Mg^{2+} concentration dependence that fit well with a Hill equation to yield a half-saturation point of $K_{1/2}$ ~ 0.6 mM and a cooperatively coefficient of n = 1.7 (Figure 2.5). These data demonstrate that the Mn^{2+} riboswitch adopts an extended SU conformation in the absence of divalents, which increasingly samples transient docked conformations upon a rise in Mg^{2+} concentration.

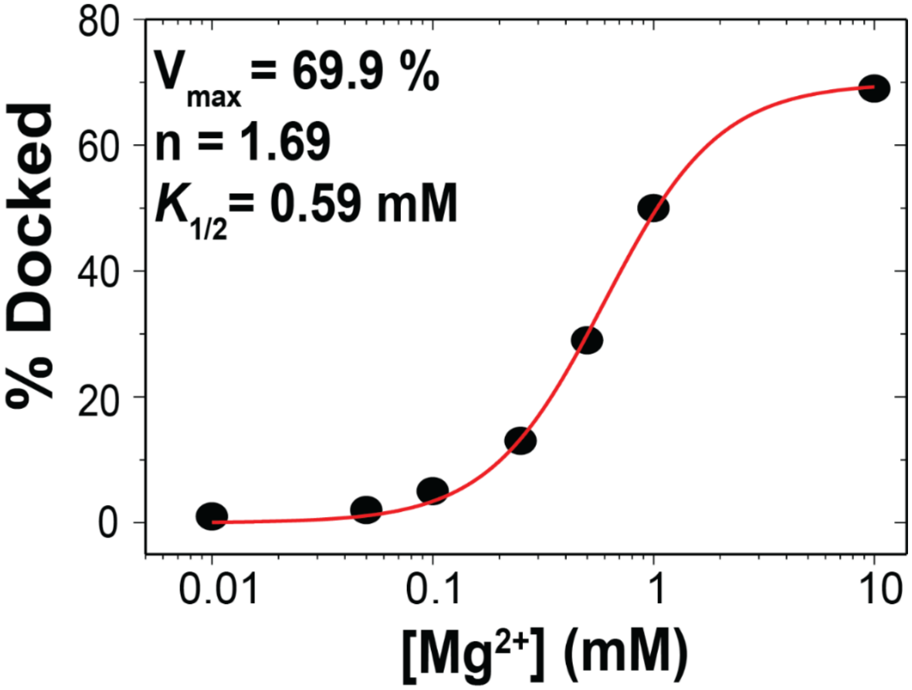


Figure 2.5 Fraction of the high-FRET state as a function of Mg^{2+} concentration, fit with a standard Hill equation (black).

At a low Mg^{2+} concentration of 0.1 mM, single-exponential kinetics were observed with a slow docking rate constant, k_{dock} ~0.56 s^{-1}, and a fast undocking rate constant, k_{undock} ~12.5 s^{-1} (Figure 2.6 and Figure 2.7). Further increasing the Mg^{2+} concentration to 1 mM resulted in the emergence of double-exponential kinetics in both k_{dock} and k_{undock}. The docking kinetics exhibited

k_{dock}^{fast} ~6.25 s^{-1} and k_{dock}^{slow} ~0.76 s^{-1}, while the undocking kinetics displayed k_{undock}^{fast} ~3.70 s^{-1} and k_{undock}^{slow} ~1.21 s^{-1} (Figure 2.6 and Figure 2.7). The TODP further shows that a majority (82 %) of traces are dynamically transitioning between the two FRET states, as highlighted by dominant 'off-diagonal' contours, while only a small fraction (~18 %) remains in the stable low-FRET state (Figure 2.3c).

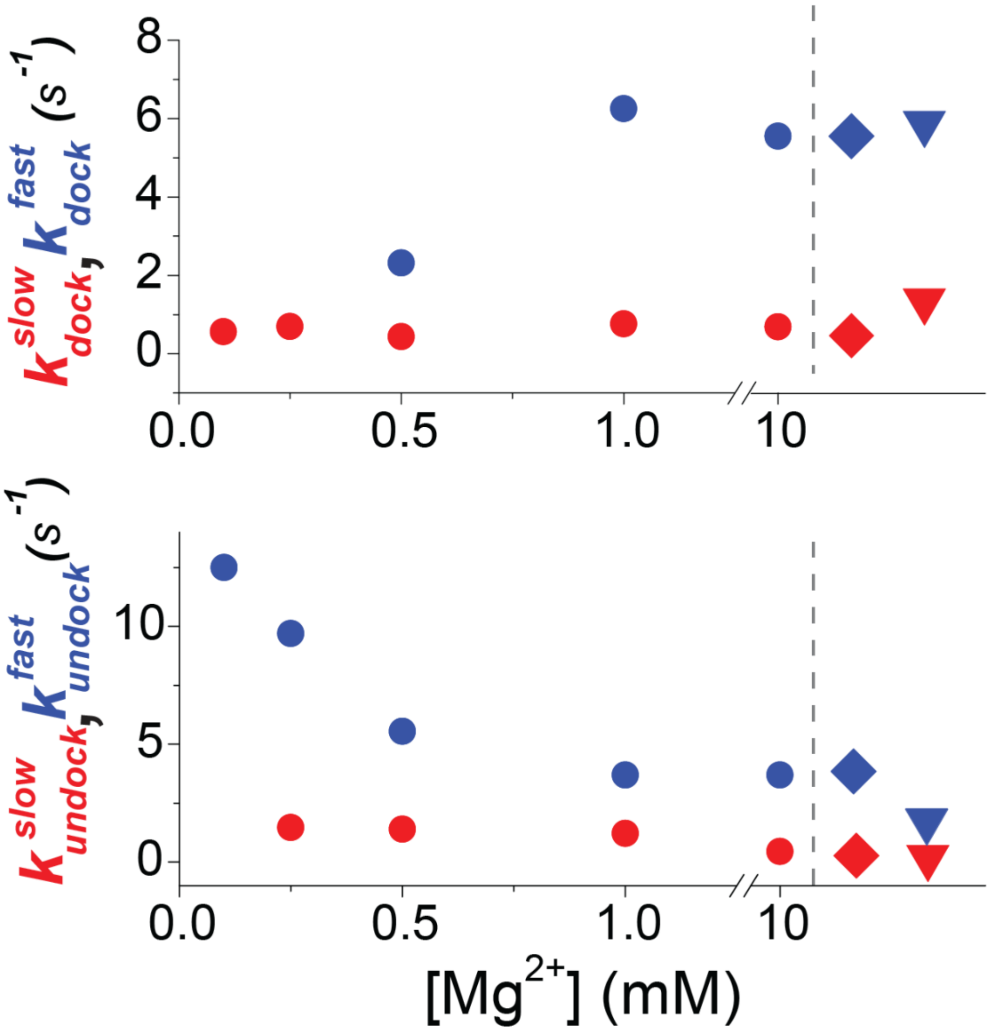


Figure 2.6 Kinetics of structural dynamics as a function of Mg^{2+} and Mn^{2+} concentration.

The diamond symbols represent rates in 1 mM Mg^{2+} and 0.1 mM Mn^{2+} while the triangle symbols represent rates in 0.1 mM Mn^{2+} alone.

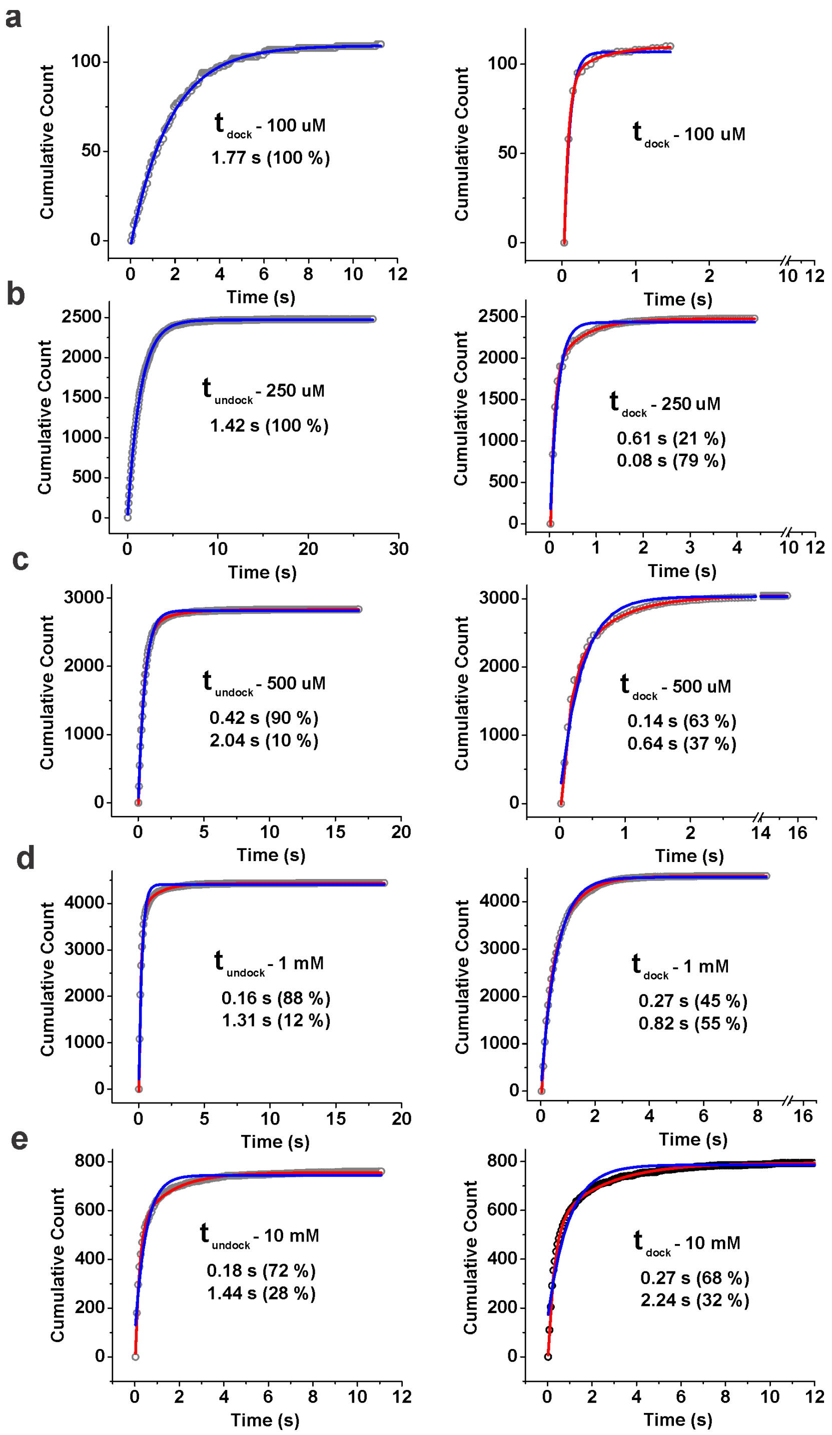


Figure 2.7 | Kinetics of WT *Xory* Mn^{2+} sensing riboswitch.

(a) Cumulative dwell-time distributions of t_{undock} and t_{dock} in the presence of 0.1 mM $MgCl_2$ fit to single (blue) and double-exponential (red) are shown. The lifetimes and amplitudes of slow and fast components are also shown. In case of double-exponential fits, fit to single exponential function is also shown for comparison. **(b-e)** Same as in (a) but in the presence of 0.25 mM, 0.5 mM, 1 mM and 10 mM $MgCl_2$, respectively.

The double-exponential kinetics arises from two distinct populations: dynamic docked (DD) and dynamic undocked (DU) traces corresponding to molecules residing largely in the docked and undocked states, respectively. (Figure 2.8a). Among the dynamic traces, ~65 % were DD while ~35 % were DU traces. As observed for other RNAs that undergo docking of two adjacent helical arms[115-121], the heterogeneity observed in the population was largely static and molecular behaviors interconverted only rarely (< 2% of traces) over the experimental timescale (5-10 min) (Figure 2.8b). Interconversion between the DU and DD behaviors was observed more readily, however, when first chelating, then reintroducing Mg^{2+} (Figure 2.8c), suggesting that they represent kinetically trapped conformations on a deeply rugged folding free energy landscape[122,123].

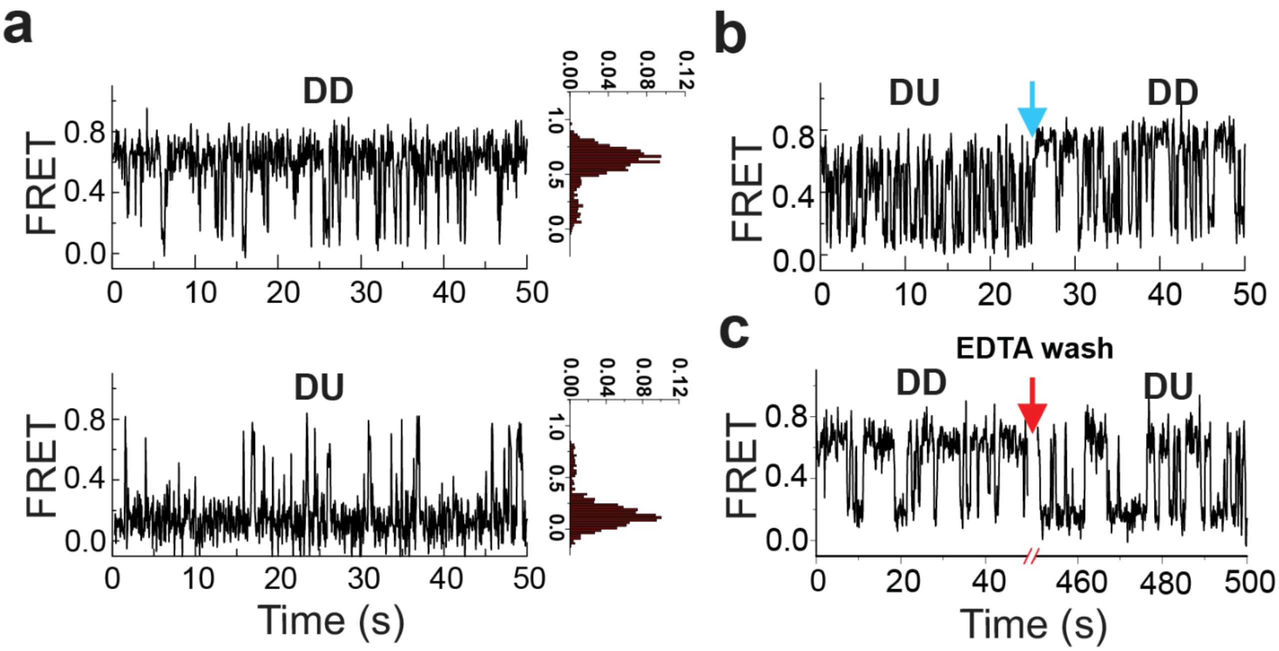


Figure 2.8 Exemplary traces with different types of dynamics in the presence of Mg^{2+}.

(a) Dynamic docked (DD, top) and dynamic undocked (DU, bottom) traces in the presence of 1 mM Mg^{2+}. FRET histograms for the individual traces are shown on the right. (b) Rare examples of interconversion between different kinetic regimes showing dynamic heterogeneity. The blue arrow indicates the time of switching between the two kinetic regimes. (c) Exemplary trace showing conversion of a DU trace to DD trace upon chelation of Mg^{2+} with EDTA (indicated by the red arrow) and reintroducing 1 mM Mg^{2+}.

2.3.2 Sub-millimolar Mn^{2+} uniquely yields a stable docked riboswitch

We next asked what specific effect Mn^{2+} has on the folding of the riboswitch. In the presence of 1 mM Mg^{2+}, addition of 0.1 mM Mn^{2+} resulted in the appearance of a unique population of stable docked (SD, ~43%) traces residing in the high-FRET state for >30 s ($k_{undock} <$ 0.03 s^{-1}) before photobleaching (Figure 2.9a). Accordingly, the FRET histogram showed two peaks with mean FRET values of 0.17±0.14 and 0.69±0.12 and an increased ~68 % population of the docked conformation (Figure 2.9a). The SD population is evident in the TODP as a new on-diagonal contour centered on the ~0.7-FRET value (Figure 2.9a). Double-exponential kinetics similar to the 1 mM Mg^{2+} alone condition were observed, with k_{dock}^{fast} = 5.55 s^{-1}, k_{dock}^{slow} = 0.46 s^{-1}, k_{undock}^{fast} = 3.84 s^{-1} and a ~5-fold slower k_{undock}^{slow} of 0.23 s^{-1}. Notably, in the presence of Mn^{2+}, most of the dynamic traces showed DD character. These data demonstrate that Mn^{2+} binding stabilizes the docked conformation while uniquely enabling an SD state. FRET histograms further showed that out of a variety of divalent metal ions tested, only Cd^{2+} is effective in promoting docked conformations over just 1 mM Mg^{2+} alone; Ni^{2+}, Co^{2+}, Sr^{2+} or Zn^{2+} had little effect (Figure 2.9b). Interestingly in the absence of Mg^{2+}, 0.1 mM Mn^{2+} alone led to the appearance of DD and SD traces with ~62 % docked population (mean FRET 0.67±0.12) (Figure 2.9c).

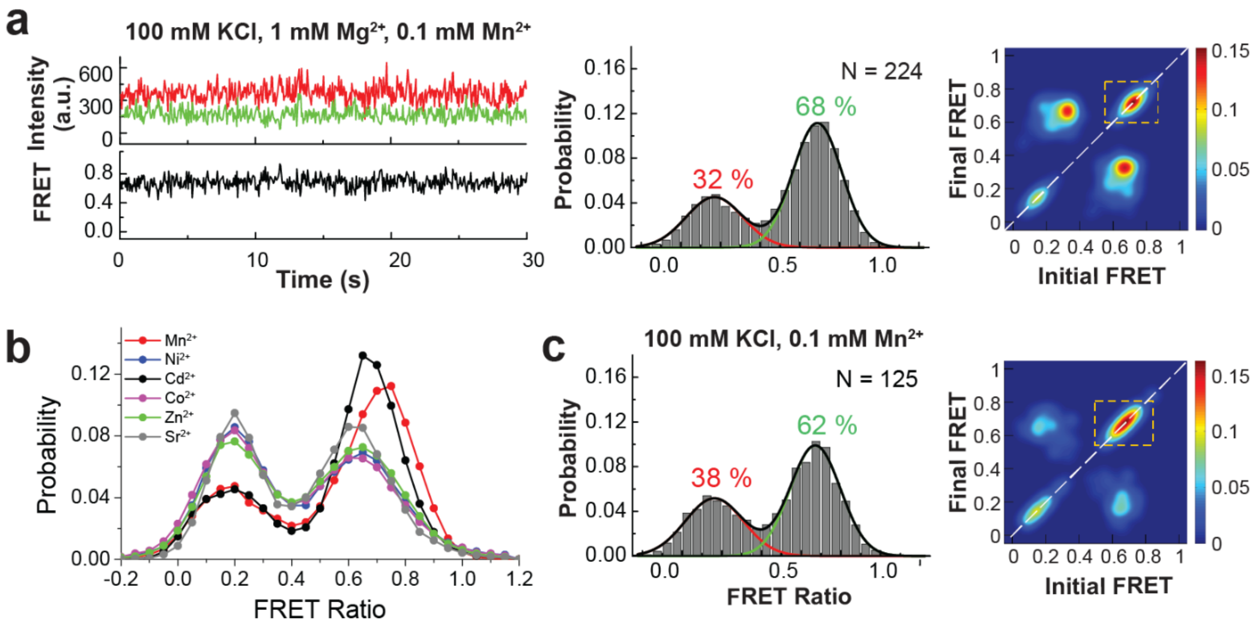


Figure 2.9 Effect of Mn^{2+} on the conformational dynamics of the riboswitch.

Representative smFRET trace, FRET histogram and TODP for the WT riboswitch in the presence of 0.1 mM Mn^{2+} and 1 mM Mg^{2+}. The stable 'static' docked (SD) conformation appears only in the presence of Mn^{2+}, as highlighted by the dashed yellow box in the TODP. (**b**) FRET histograms for the WT riboswitch at 0.1 mM concentration of different transition metals in the presence of 1 mM Mg^{2+}. Over 150 molecules were analyzed for each histogram. (**c**) FRET histogram and TODP with SD population highlighted by dashed yellow box for the WT riboswitch in the presence of 0.1 mM Mn^{2+} alone.

To probe how the kinetically distinct SU, DU and DD traces respond to Mn^{2+}, we performed ligand jump experiments and mapped their transition from one conformation to another (Figure 2.10). We observed the same set of molecules (N=160) at 1 mM Mg^{2+}, both before and after the addition of Mn^{2+} for equal time window. With only 1 mM Mg^{2+}, we observed three groups population with SU, DU and DD conformations by individually counting them. In this condition, most fraction of molecules (53%) went from SU to DD and a significant fraction (34%) to DU. A small fraction (14%) didn't undergo any conformation change and stay as SU. Interestingly, none of the population went to stable docked (SD) conformation with only 1 mM Mg^{2+}, supporting earlier observations from histograms and TODPs (Figure 2.4 & Figure 2.9). Upon addition of 0.1 mM Mn^{2+} to the same molecules, we found that all three populations respond to Mn^{2+} and were capable of forming the SD population. In particular, the SU traces converted into DU (23%), DD (23%), and SD (55%) conformations with similar probabilities only upon the addition of Mn^{2+} to the solution, suggesting that they are correctly folded with metal sensing sites poised to bind Mn^{2+} ions selectively. By comparison, most DD traces converted into SD traces whereas DU traces adopted DD and SD behavior upon Mn^{2+} addition. In the population group of DU and DD, 52% and 85% of their population converted to SD respectively. Only a small fraction of traces showed no response at such low concentration Mn^{2+} suggesting that they may be misfolded.

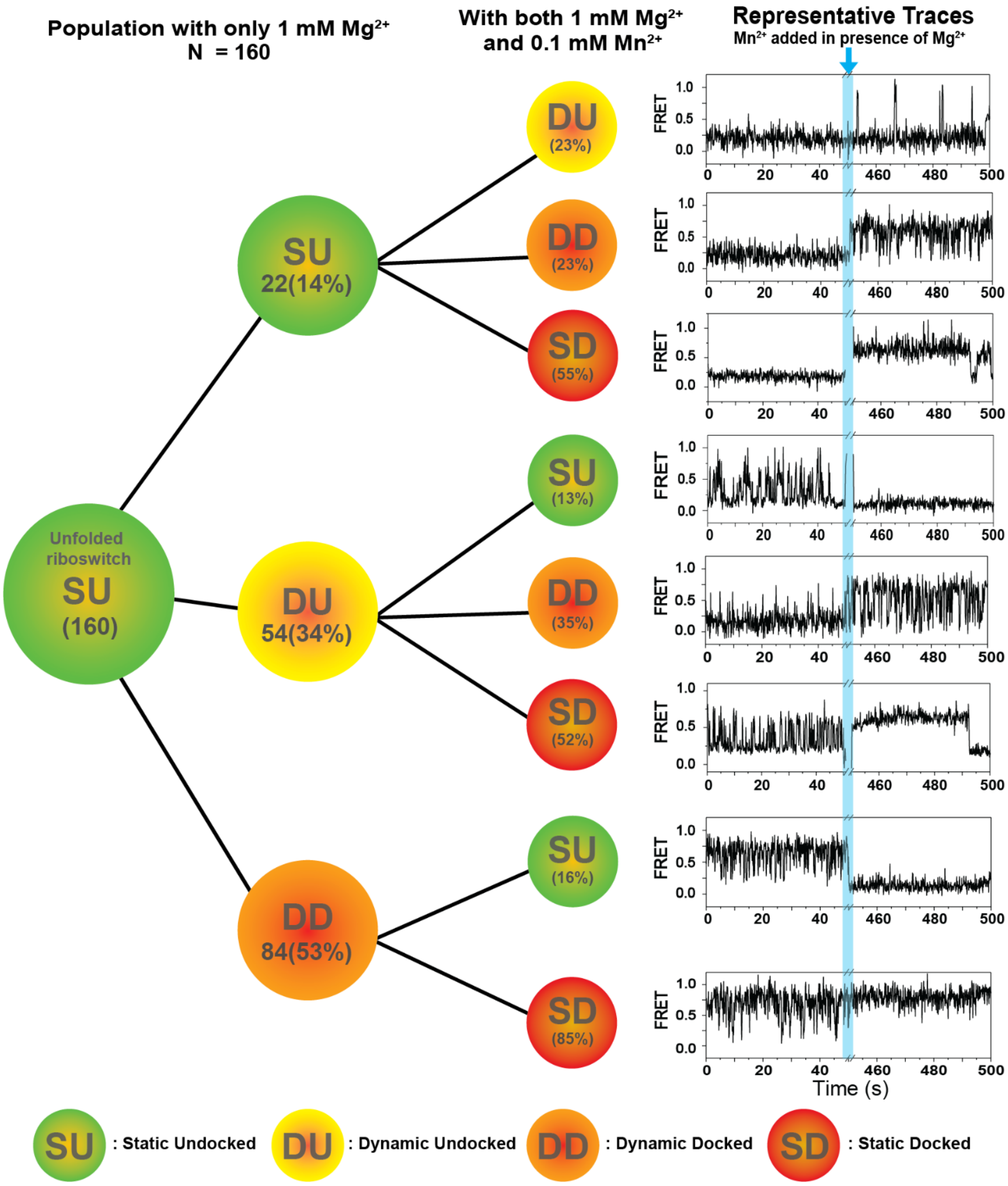


Figure 2.10 Distribution of population transitioning between different conformations (SU, DU, DD and SD) without and with the addition of 0.1 mM Mn^{2+} ions and their representative traces

2.3.3 The *yybP-ykoY* riboswitch can discriminate between similar transition metal ions

The selectivity of the *Xory* riboswitch for Mn^{2+} over Mg^{2+} arises in part from the inner-sphere contact with A48(N7), suggesting that other soft transition metal ions may also be recognized[111]. To test this hypothesis further, we probed the effects of different divalent metal ions on the conformation of the WT riboswitch, at 0.1 mM concentration alone or in the presence of 1 mM Mg^{2+}. FRET histograms showed that out of all the different metal ions tested, Cd^{2+} is most effective in promoting docked conformations (Figure 2.9b). To further confirm the interaction of different divalent ions to the metal ion binding core of the riboswitch, we looked at the transition dynamics and stable docking population by plotting their TODP in comparison with the histogram for each condition (Figure 2.11). In the presence of 1 mM Mg^{2+}, addition of 0.1 mM Cd^{2+} resulted in ~65 % of the high-FRET docked conformation, comparable to the docked population upon addition of 0.1 mM Mn^{2+}. Examination of individual smFRET traces as well as the TODP showed that this is due to a large fraction of SD traces (Figure 2.11), in agreement with the tight binding of Cd^{2+} to the *yybP-ykoY* riboswitch shown recently[111]. Among the other metals tested, Ni^{2+}, Co^{2+}, Sr^{2+} or Zn^{2+} had little effect on promoting the folded conformations of the riboswitch under these conditions. Interestingly, in the absence of Mg^{2+}, while 0.1 mM Mn^{2+} alone led to the appearance of DD and SD traces with ~62 % docked population (mean FRET 0.67±0.12) (Figure 2.9b), 0.1 mM of Ni^{2+}, Co^{2+}, Sr^{2+} or Zn^{2+} did not affect SU traces and Cd^{2+} had only a small effect in promoting DD traces (Figure 2.11). This suggests that, while Mg^{2+} and Mn^{2+} may both bind at $M_{A,Mg}$, Cd^{2+} may be more specific to the $M_{B,Mn}$ site, although the affinity of Cd^{2+} as compared to Mn^{2+} may be much lower due to the difference in their ionic radii. These results suggest that while the *Xory* riboswitch has some degree of plasticity in recognizing ligands. In a background of Mg^{2+}, it preferentially recognizes Mn^{2+} and and – to a lesser extent – Cd^{2+}, while effectively discriminating against similar divalent transition metal ions.

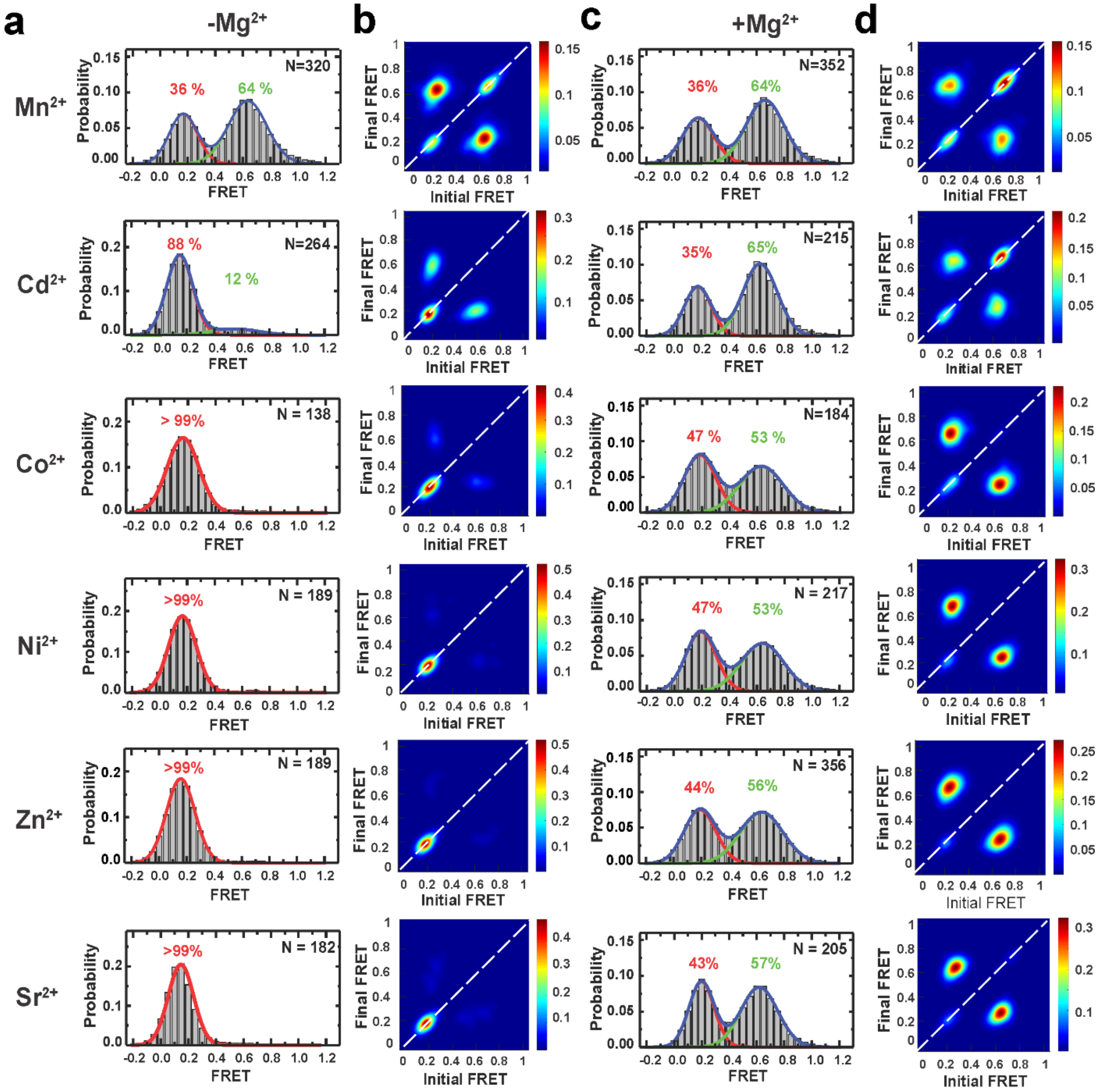


Figure 2.11 Effect of different metal ions on the WT *Xory* riboswitch.

(a) FRET histograms with equilibrium distribution of docked and undocked conformations of the WT riboswitch at 0.1 mM concentration of different transition metal ions alone (i.e., in the absence of 1 mM $MgCl_2$). Gaussian fits are shown as green and red curves, while the cumulative fit is shown in blue. The percentage (%) of docked and undocked conformations in each condition are indicated. N represents number of molecules for each condition. **(b)** TODPs corresponding to the histograms in (a). **(c)** FRET histograms and **(d)** TODPs at 0.1 mM concentration of different metal ions in the presence of 1 mM $MgCl_2$. For the different metal ions tested, except Mn^{2+} and Cd^{2+}, almost all the traces remained in SU conformation, with >95 % in the low-FRET (~0.1) undocked conformation, in the absence of Mg^{2+}.

2.3.4 Mutation of A48 results in complete loss of SD conformation

The highly conserved discriminator base A48 in L3 is positioned to confer Mn^{2+} specificity via its N7 and also helps maintain L3 in a stacked conformation. We therefore tested the effect of a single A48U mutation on folding and Mn^{2+} sensing of the riboswitch. Similar to the WT, at 100 mM KCl without divalents, smFRET traces of the mutant riboswitch showed the SU population with a mean FRET value of 0.11±0.12 (Figure 2.12a). At 1 mM Mg^{2+}, we observed dynamic traces with excursions into the docked higher FRET states, and the FRET histogram showed a major (64 %) 0.14±0.11 low-FRET peak and a minor (36 %) broad 0.52±0.21 mid-FRET peak that corresponds to more extended docked conformations (Figure 2.12b). Of note, 100% of the dynamic mutant traces showed DU character under this condition (Figure 2.12b). As a result, we observed fast single-exponential k_{undock} ~6.67 s^{-1} while k_{dock} was double-exponential with a major k_{dock}^{fast} = 3.70 s^{-1} (95 %) and a minor k_{dock}^{slow} = 0.46 s^{-1} (5 %) (Figure 2.13a). The TODPs show SU behavior in the absence of divalents, whereas both SU and DU behaviors are observed at 1 mM Mg^{2+} (Figure 2.12a & b).

Next, we asked whether the mutant riboswitch can still respond to Mn^{2+}. In the presence of 1 mM Mg^{2+} and 0.1 mM Mn^{2+}, most smFRET traces showed DD behavior, while only a small fraction remained in the SU state, similar to WT (Figure 2.12c). FRET histograms show a similar ~0.14±0.13 (57 %) low-FRET undocked state but display a docked state now with a higher mean-FRET value of 0.63±0.15 and a larger ~43 % population, compared to the 1 mM Mg^{2+} only condition. The kinetics under these conditions were double-exponential with a k_{dock}^{fast} = 6.67 s^{-1} (89 %), k_{dock}^{slow} = 0.64 s^{-1} (11 %), k_{undock}^{fast} = 3.10 s^{-1} (94 %), k_{undock}^{slow} = 0.88 s^{-1} (6 %) (Figure 2.13b). Importantly, the mutation caused a striking loss of SD traces, as evident from the complete absence of the on-diagonal ~0.7 FRET contour in the TODP (Figure 2.12c), in stark contrast to the WT. In addition, the mutant interestingly loses the WT's ability to sample the docked conformations in the presence of 0.1 mM Mn^{2+} alone (Figure 2.12d); almost all smFRET traces at 0.1 mM Mn^{2+} are in the SU conformation, with the histogram showing a major single peak around 0.19±0.13 FRET, further corroborated by a single on-diagonal ~0.2 FRET contour in the TODP (Figure 2.12d). These data demonstrate that the conserved discriminator A48 is essential for Mn^{2+} inducing a stable docked riboswitch conformation. The presence of Mg^{2+} partially rescues the loss of Mn^{2+} sensing by the mutant yet does not restore its ability to form the SD conformation. Possibly this reflects the ability of Mg^{2+} to bind at M_B in A48U, though with weaker affinity, via the uridine O4, as suggested in the A41U *L. lactis* structure[106].

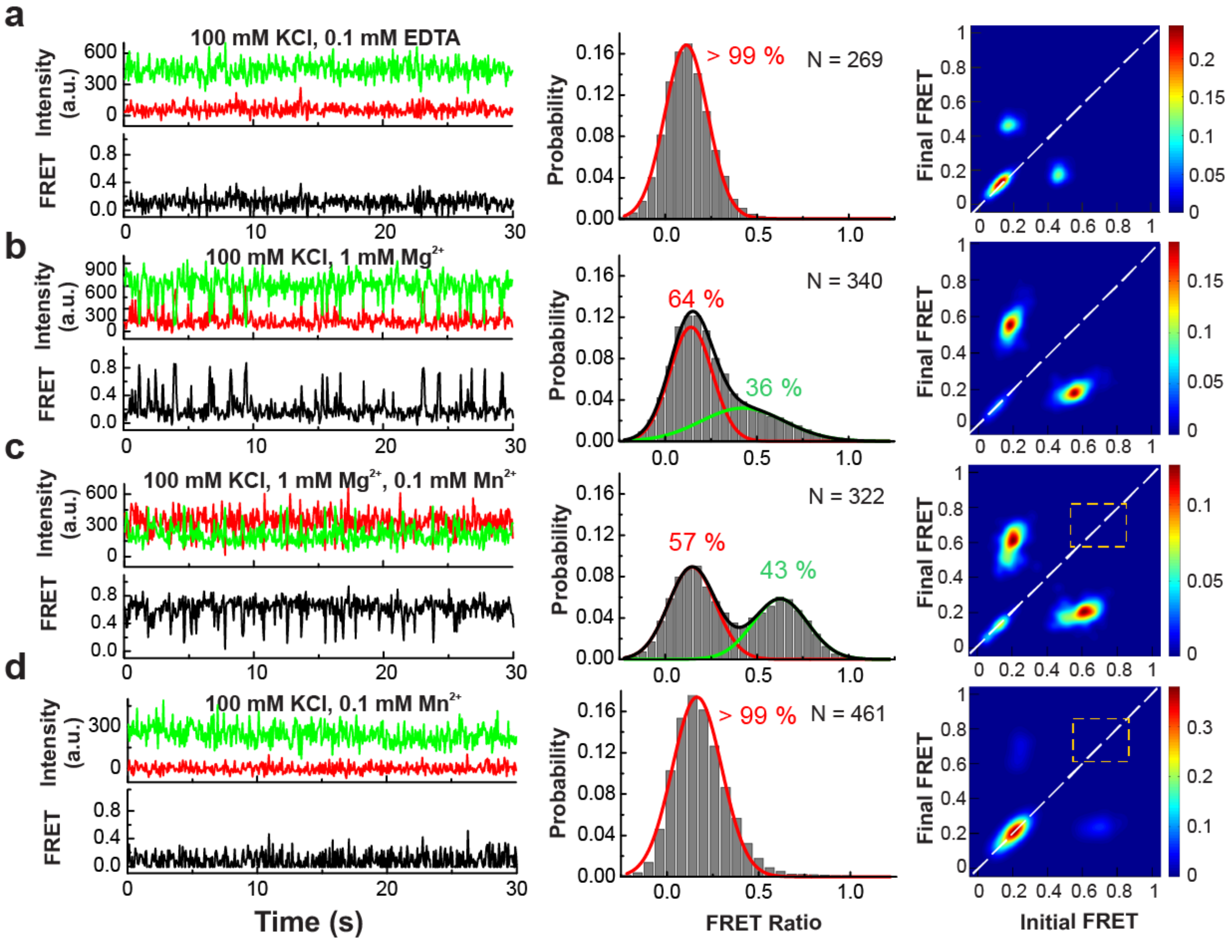


Figure 2.12 Mutation of the Mn^{2+} sensing A48 leads to loss of the SD conformation.

Representative smFRET traces, FRET histograms and TODPs for the A48U mutant riboswitch under: (**a**) no divalents (+0.1 mM EDTA), (**b**) 1 mM Mg^{2+} alone, (**c**) 1 mM Mg^{2+} and 0.1 mM Mn^{2+}, and (**d**) 0.1 mM Mn^{2+} alone. The dashed yellow box highlights the absence of SD conformations for the A48U riboswitch in the presence of Mn^{2+}, in contrast to the WT riboswitch.

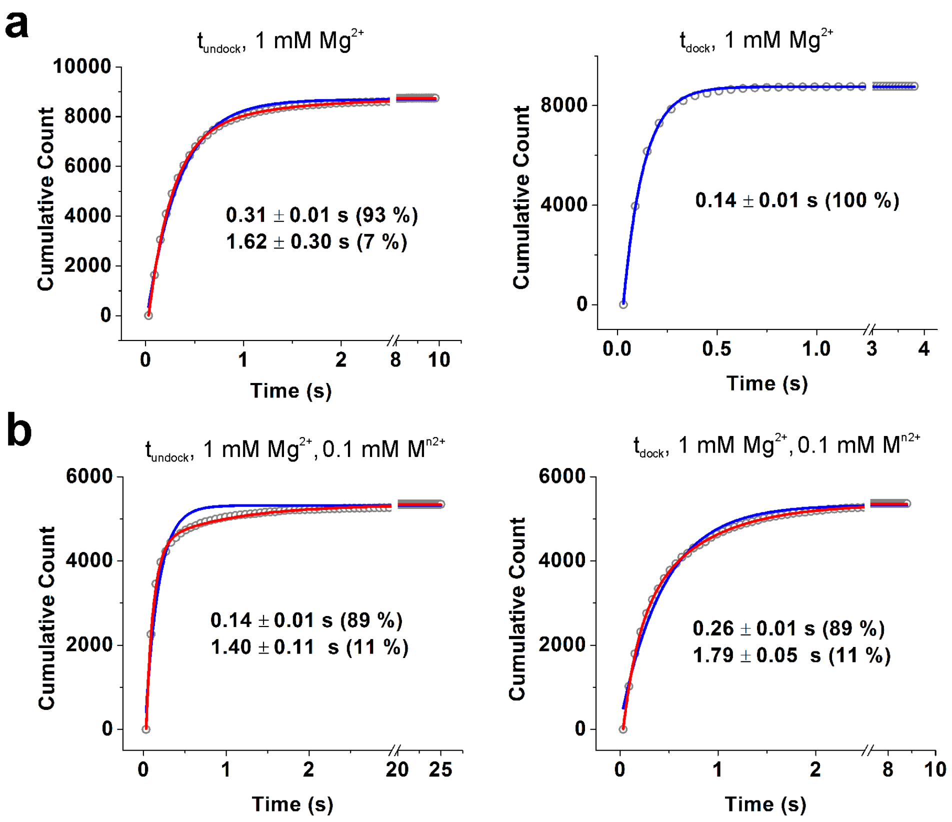


Figure 2.13 Kinetics of the *Xory* riboswitch A48U mutant.

(a) Cumulative dwell-time distributions of t_{undock} and t_{dock} in the presence of 1 mM $MgCl_2$ fit with single (blue) and double-exponential (red) functions. The lifetimes and amplitudes of slow and fast components are also shown. In the case of double-exponential fits, fit to single exponential is also shown for comparison. (b) Same as in (a) but in the presence of 1 mM $MgCl_2$ and 0.1 mM $MnCl_2$.

2.3.5 Mn^{2+} binding to the riboswitch enhances the stability of switch helix P1.1

The proposed mechanism of transcription antitermination of the *yybP-ykoY* riboswitch, supported by our structural and dynamic data, involves Mn^{2+}-mediated stabilization of the inherently weak P1.1 switch helix. To more directly test this prevailing hypothesis, we probed the accessibility of the 3'-half of the P1.1 helix (last 12-nt) *in trans* by a complementary Cy5-labeled DNA oligonucleotide probe using our established Single Molecule Kinetic Analysis of RNA Transient Structure (SiM-KARTS) assay (Figure 2.14a)[65]. In the current configuration, SiM-KARTS mimics the strand invasion by the downstream terminator sequence and reports on the stability of P1.1 helix in WT and A418U mutant riboswitch at the single-molecule level (Figure 2.14b & c).

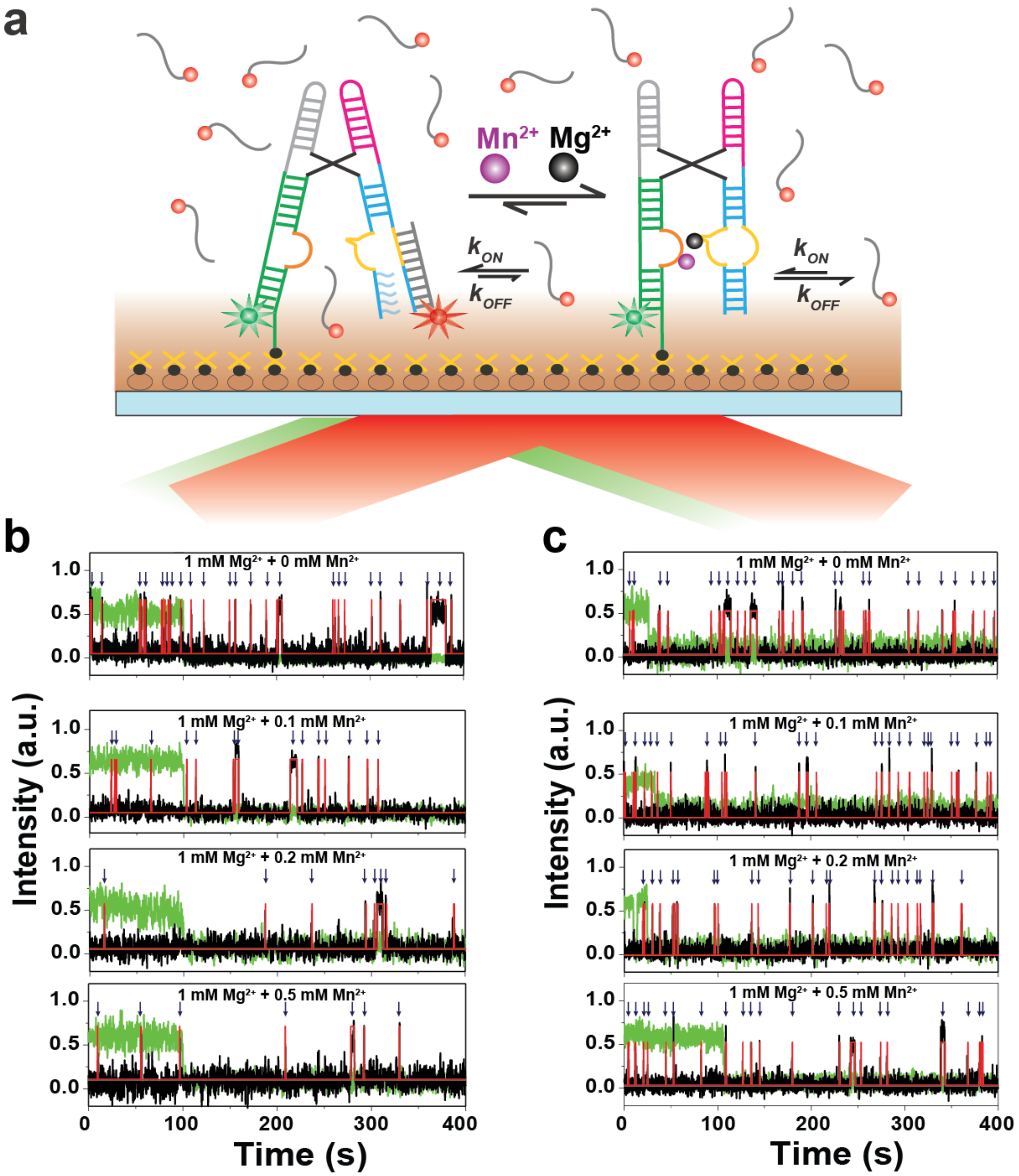


Figure 2.14 Accessibility of switch helix P1.1 probed using a SiM-KARTS assay.

(a) Schematic of our SiM-KARTS assay to probe the accessibility of switch helix P1.1 in the presence and absence of Mn^{2+}. (b) and (c) Representative SiM-KARTS traces at different Mn^{2+} concentrations show multiple binding events of the Cy5-labeled DNA oligonucleotide probe for WT and A48U mutant riboswitch.

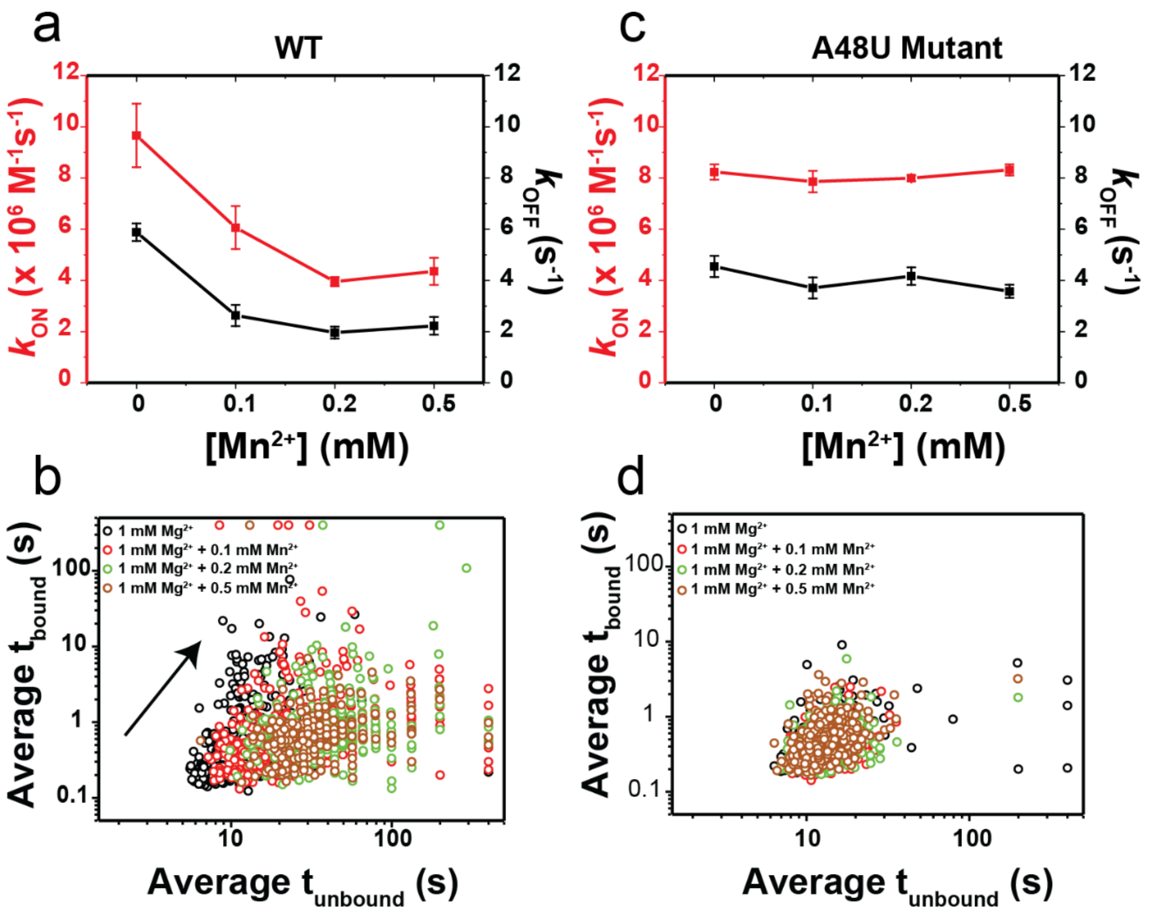


Figure 2.15 Rate constants and average dwell times for SiM-KARTS assay.

(a) The association (k_{on}, red) and dissociation (k_{off}, black) rate constants and (b) scatter plot of average bound and unbound dwell times for binding of probe to P1.1 as a function of Mn^{2+} concentration for the WT riboswitch. (c) The association (k_{on}, red) and dissociation (k_{off}, black) rate constants and (b) scatter plot of average bound and unbound dwell times for binding of probe to P1.1 as a function of Mn^{2+} concentration for the A48U mutant riboswitch. Over 200 traces were analyzed for each condition to obtain reliable kinetic rate constants.

In the presence of 10 nM Cy5-labeled oligonucleotide and 1 mM Mg^{2+}, repeated short binding events were observed with a probe binding rate constant k_{on} of 9.67±1.24 x 10^6 M^{-1} s^{-1} and a dissociation rate constant k_{off} of 5.88±0.35 s^{-1} for the WT riboswitch (Figure 2.15a). Addition of increasing concentrations of Mn^{2+} resulted in a decrease in the frequency of binding events, with a k_{on} value significantly reduced by ~2.5-fold in the presence of 200 µM Mn^{2+} (Figure 2.15a). These data report on a decreasing accessibility of the 3'-segment of the P1.1 helix and thus a more stable P1.1 duplex in the presence of Mn^{2+}. Interestingly, the k_{off} value also decreases with the addition of Mn^{2+} (Figure 2.15a & b), consistent with the notion that the RNA can still sense Mn^{2+} via interaction of the free 5'-half of P1.1 – extending into the L1 loop – with loop L3, resulting in a stable L1 and P1.2 that can stack on top of the RNA-DNA hybrid. Remarkably, the effect of Mn^{2+} is completely lost in the A48U mutant, where the k_{on} and k_{off} values remain essentially unchanged at $8.24 \pm 0.30 \times 10^6$ M^{-1} s^{-1} and 4.54 ± 0.41 s^{-1}, respectively, even at high (500 μM) Mn^{2+} concentration, showing no change in the frequency of individual binding events (Figure 2.15c & Figure 2.16c & d). These data demonstrate that the stability of P1.1 is enhanced once the riboswitch encounters Mn^{2+}, an effect lost upon mutation of the Mn^{2+} sensing adenosine. These results thus provide a direct evidence for the prevailing mechanistic hypothesis of riboswitching by the *yybP-ykoY* RNA, wherein Mn^{2+} binding stabilizes the antiterminating P1.1 switch helix through long-range signal transduction from the metal sensing core, thereby preventing formation of the competing terminator hairpin.

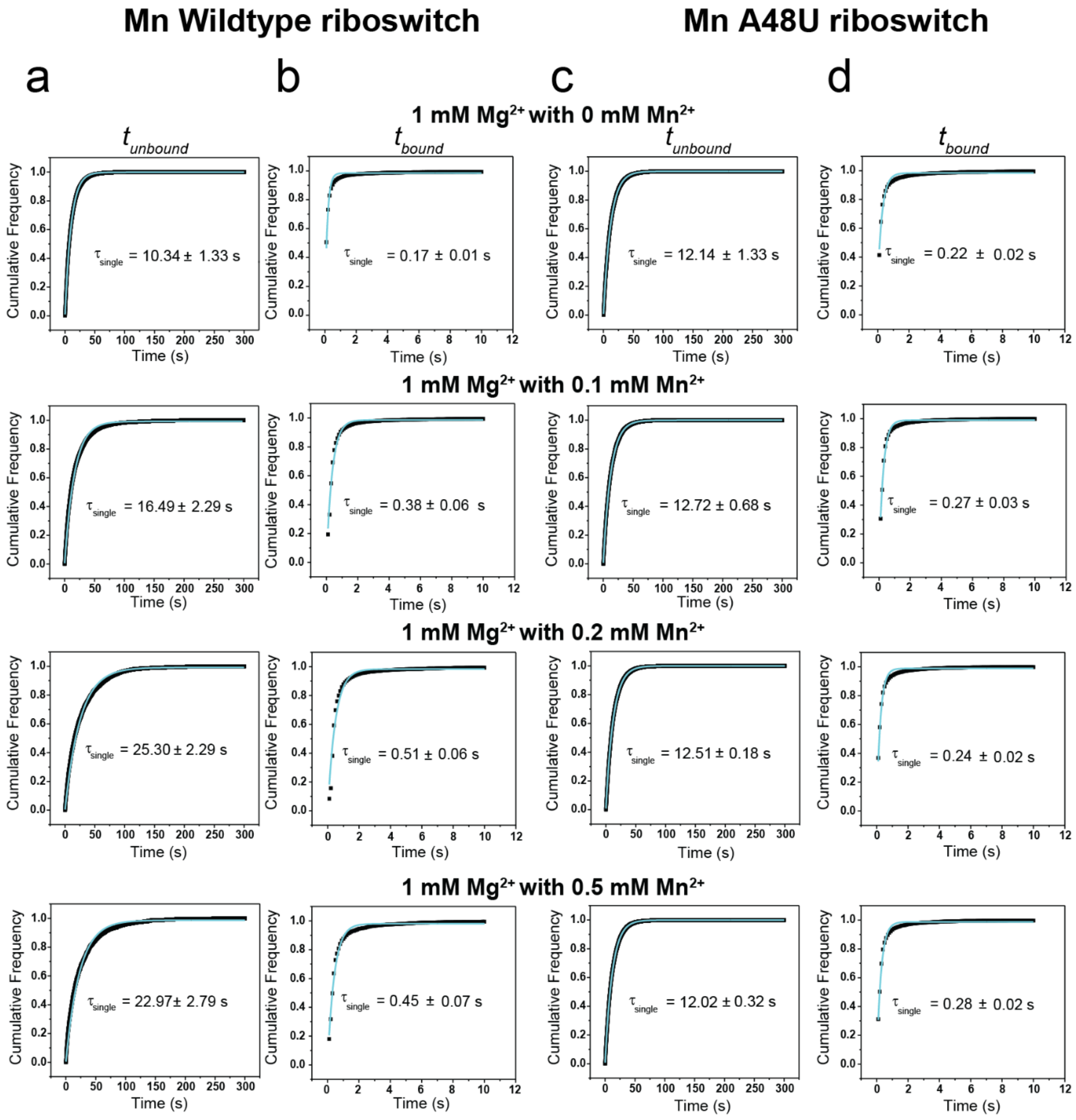


Figure 2.16 Cumulative dwell-time distribution plots for SiM-KARTS data for the WT and A48U mutant riboswitch.

Distribution plots of (a) & (c) $t_{unbound}$ and (b) & (d) t_{bound} for WT and A48U mutant fitted to single-exponential functions to obtain k_{on} and k_{off} in the presence of 1 mM Mg^{2+} with 0 mM Mn^{2+}, 0.1 mM Mn^{2+}, 0.2 mM Mn^{2+}, and 0.5 mM Mn^{2+}. The single-exponential fits to the data are shown in cyan along with the individual unbound and bound lifetimes.

2.4 Discussion

Using a combination of X-ray crystallography, MD simulations and smFRET, our work here sheds light on how the binding of a single Mn^{2+} ion at a distant site can modulate the stability of the switch helix through a conformational cascade that couples local with global structural dynamics. While previous structures revealed the local Mn^{2+} recognition mechanism of *yybP-ykoY* RNA[106,111], the structures reported here underscore the conformational plasticity around the metal binding core. Importantly, the structural differences observed between the two conformers in the crystal lead us to posit that conformer 2 with L3 partially unfolded around $M_{B,Mn}$ is likely an intermediate between the fully Mn^{2+}-bound state and the empty $M_{B,Mn}$ state observed in the *E. coli* Mn^{2+}-free structures[106]. We hypothesize that the loss of phosphate contacts from C51, U52, and G9 to Mn^{2+} in conformer 2 reflects their relative weakness as supported by MD simulations and likely represents a discrete step in Mn^{2+} binding/unbinding. Conversely, the variation in $M_{B,Mn}$ site coordination we observe could represent inherent flexibility at this site, allowing it to recognize Mn^{2+} even if the metal ion is not fully dehydrated. This flexibility is also seen in the recent Cd^{2+}-bound structures[111]. The MD simulations data (Appendix Figure 6.2) further show that the SRL-like motif in L1 is inherently unstable in the L1-L3 undocked state, but is stabilized by the tertiary A-minor interaction of A10 with P3.1. Similarly, L1-L3 docking stabilizes L3 in a stacked conformation by direct stacking of A10 on top of it. Thus, a stable conformation of L1 and therefore P1.1 is intimately linked to the stability of L3 and the tertiary A-minor interaction. We suggest that the SRL-like motif acts as a molecular switch that links the Mn^{2+}-dependent increase in docking observed by smFRET to stabilization of P1.1.

Using smFRET, we further identify a globally undocked conformation at physiological Mg^{2+} concentration that remains in dynamic equilibrium with the compact docked conformations. In fact, the observation of pairs of stably and dynamically docked and undocked states with double-exponential docking and undocking kinetics supports the existence of at least two undocked and docked dynamic (in addition to the static SU and SD) conformations, with similar FRET values and thus global folds, but likely different local conformations around the ligand sensing core. These conformations are possibly related to conformers 1 and 2 in the crystal structure. We further find that SD conformations are formed only in the presence of sub-millimolar Mn^{2+} or the related soft transition metal Cd^{2+} and lost upon a single mutation of the invariant Mn^{2+}-sensing adenine

A48. While our work was in preparation, another smFRET study on the folding of a different *Llac* Mn^{2+} riboswitch became available. Using a similar labeling strategy, this *Llac* riboswitch showed two FRET states similar to the *Xory* riboswitch[124]. Interestingly, the kinetics of conformational dynamics for the *Llac* riboswitch were largely single-exponential, suggesting a two-state folding behavior, that agrees with a single conformation of the ligand-bound structure. This is in contrast to the *Xory* riboswitch studied here, which shows double-exponential kinetics, consistent with the two distinct conformations of the ligand-bound structure seen in our crystal structures. Furthermore, while the *Llac* riboswitch shows a single-exponential k_{undock} that remains constant with increasing Mg^{2+} concentrations, the k_{undock} for *Xory* riboswitch shows double-exponential behavior, where the fast k_{undock} decreases with increasing Mg^{2+}. In addition, compared to the *Llac* riboswitch, the *Xory* riboswitch shows higher cooperativity in Mg^{2+} binding, suggesting that the $M_{B,Mn}$ site of *Xory* riboswitch can bind Mg^{2+}, albeit weakly, leading to the appearance of the dynamic docked (DD) traces under high Mg^{2+} concentrations, and resulting in the observed decrease of k_{undock}. In addition, the *Xory* riboswitch can adopt SD conformations in the presence of Mn^{2+}, which are not observed in the *Llac* riboswitch. Such species-specific differences in the Mg^{2+} and Mn^{2+} mediated folding of the *yybP-ykoY* riboswitches likely enable bacteria to elicit a fine-tuned response to changes in their intracellular metal ion concentrations.

Our structural snapshots, MD simulations and smFRET data support the long-range structural signal transduction model (Figure 2.17**)**. In the absence of any divalent ions, the *Xory* riboswitch adopts an extended conformation in which the two legs are distal and rarely interact, as shown by smFRET. Under these conditions, the metal-binding sites are likely unstructured and their interaction only transiently formed by L1-L3 contacts, as suggested by rare smFRET transitions, MD simulations, as well as previous ligand-free structures[106]. MD simulations in particular show that the A-minor interaction formed by A10 of L1 loop with the GC base pair of stem P3.1 can at least transiently form even in the absence of divalents. The formation of this A-minor interaction helps pre-organize the $M_{A,Mg}$ binding site. Under physiological (1 mM) concentrations of Mg^{2+}, this ion binds first at the $M_{A,Mg}$ site, allowing the riboswitch to sample dynamic folded conformations where the two legs are brought together via metal-mediated interaction between loops L1 and L3. These dynamic folded states then allow for local interactions mediated by L3, forming a pocket of high negative charge potential poised to sense Mn^{2+}. Mg^{2+} binding at $M_{A,Mg}$ therefore facilitates binding of Mn^{2+} at $M_{B,Mn}$ in cooperative fashion, as shown by smFRET at the global level and further visualized by atomistic MD simulations. Capture of Mn^{2+} at site $M_{B,Mn}$ then acts as the final linchpin to hold the two legs together most stably. This compact docked conformation with a SRL-like fold of L1 and continuous coaxial stacking between P1.2, L1 and P1.1 stabilizes the remote P1.1 switch helix, as shown by our SiM-KARTS assay and MD simulations, that prevents the strand invasion required to form the terminator hairpin, instead promoting transcriptional read-through.

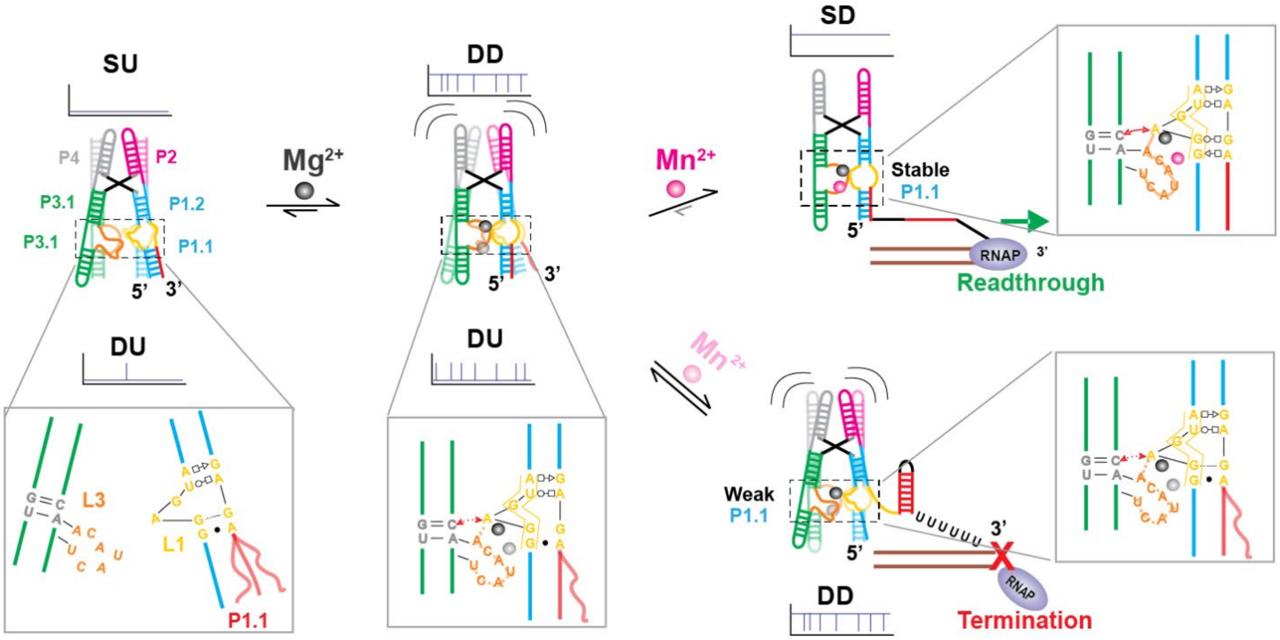


Figure 2.17 Local-to-global signal transduction pathway in the core of the *Xory* Mn^{2+} sensing riboswitch.

The *Xory* riboswitch exists in an 'X-shaped' extended SU conformation in the absence of Mg^{2+}, with very rare transitions into docked 'H-shaped' conformations. In the SU state, the legs are far apart in the absence of a stable A-minor tertiary interaction and loops L3 and L1 are less structured under these conditions. In the presence of physiological Mg^{2+} concentrations (mM), the riboswitch samples docked conformations with distinct kinetics, indicative of differences in L3 stacking as observed in our crystal structures. L1 adopts an S-turn conformation (as indicated by the yellow box in the inset) and L3 is partially stacked under these conditions, aided by the presence of Mg^{2+} ions in $M_{A,Mg}$ and transiently in $M_{B,Mn}$. In the presence of sufficient Mn^{2+}, binding of the metal ion by A48 reinforces L3 stacking and stability of the A-minor interaction. This maintains L1 in a rigid non-canonical SRL-like conformation that enables co-axial stacking between P1.2 and the switch helix P1.1. In turn, this cascade of interactions stabilizes P1.1, thereby preventing strand invasion to favor formation of the anti-terminator required for transcriptional read-through.

Contrasting with the Mn^{2+} riboswitch, the NiCo riboswitches, which cooperatively sense the transition metals Ni^{2+} or Co^{2+}, resemble the overall 'H' shaped architecture of the Mn^{2+} riboswitch but with a tight 4WJ that lacks an analogous tertiary docking interface. Accordingly, the NiCo riboswitch appears to acts through a distinct mechanism utilizing four bound metal ions to weave together a network of interactions between the interhelical residues that stabilize the 4WJ and so prevent formation of a terminator via strand invasion[125]. Similarly, the Mg^{2+} sensing M-box riboswitch adopts an architecture with three parallel co-axially stacked helices that are brought together by the binding of six Mg^{2+} ions, leading to the formation of a terminator[126]. Therefore, both the NiCo and Mg^{2+} sensing riboswitches achieve gene regulation by employing multiple metal ions that directly interact with and stabilize the switch helix. In contrast, the Mn^{2+} riboswitch is unique in requiring only a single metal ion that does not make any direct contacts with switch helix P1.1. By combining high-resolution structural information with insights into both atomistic and global dynamics, our study outlines the mechanism by which a ligand as small as a single divalent metal ion couples local with global structure to give Mn^{2+} the ability to influence RNA folding and fine-tune gene expression by cooperating with a high background of Mg^{2+}. The general lessons revealed here of how a ligand-binding signal can be transduced across an RNA are likely to become a recurring theme among riboswitches where the ligand represents a distal structural linchpin for the switch helix.

2.5 Materials and Methods

2.5.1 RNA preparation and crystallization for x-ray crystallography

RNA was cloned, transcribed, purified, refolded with 2.5 mM Mn^{2+}, and screened for crystallization as previously described[106]. The *X. oryzae* aptamer domain sequence was modified to improve the chances of crystallization, including replacing the terminal loops in variable regions with GAAA tetraloops and adding a GG at the beginning of the sequence to increase T7 RNA polymerase efficiency[127]. We also removed a single unconserved U flip-out near the base of P1.1. In this construct, a native CA dinucleotide at the four-way junction was omitted from the sequence (Figure 2.1), possibly aiding crystallization. Initial crystal hits were obtained using the Nucleic Acid Mini-Screen (Hampton Research). This RNA crystallized at 0.2 mM in 10% MPD, 40 mM Na cacodylate (pH 7), 12 mM spermine tetrahydrochloride, 80 mM $SrCl_2$, and 20 mM $MgCl_2$ after 2-3 months at 18 °C.

2.5.2 RNA preparation for single-molecule FRET

The RNA for smFRET studies was annealed from two synthetic RNA oligonucleotides, ordered with the indicated modifications from IDT. The native CA dinucleotide omitted for crystallization was present in the smFRET RNA. Oligonucleotide 1 has a 5'-Cy5 and oligonucleotide 2 has 5'-Cy3 and 3' biotin-TEG. In this two-oligonucleotide design of the smFRET construct, the donor fluorophore Cy3 and biotin-TEG were placed on the 5' and 3' ends of stem P1.1 while the acceptor fluorophore Cy5 is attached to the 5' end of P3.2, positioning the fluorophores on either end of the distal arms that dock. For specifically probing the docking interaction, we extended the stem P1.1 to prevent end fraying. The biotin on the 3'-end of P1.1 was used for immobilization of the RNA onto a microscope quartz slide and subsequent prism-based total internal reflection fluorescence microscopy (Figure 2.2b). Optimal refolding of the RNA was achieved with a mixture of 1 µM oligonucleotide 1 and 1.5 µM oligonucleotide 2 in 20 mM HEPES pH 7.0, 50 mM KCl. In PCR tubes, samples were heated to 70 °C for 2 min, then 2 mM $MgCl_2$ was added, samples were allowed to cool to RT for 10 min, then put on ice until further use. To check assembly, ~1 pmol of each oligonucleotide (1-2 µL) was added to 10 µL running buffer supplemented with 10% glycerol and loaded onto a non-denaturing 8% polyacrylamide gel run at 4 °C. A buffer of 0.5x TBE supplemented with 2 mM $Mg(OAc)_2$ and 50 mM KCl was used in the gel and as the running buffer.

Oligo 1:
/5'Cy3/AUCCUUGGGGAGUAGCCUGCUUUCUUCGGAAAGCGCCUGUAUCAACAUAC
UCGGCUA
Oligo 2:
/5'Cy5/UAGCCGUGGUGCAGGCAACGGCGAAAGCCGUCUGGCGAGACCAGGGAU/**Bio
TEG**/3'

2.5.3 smFRET data acquisition and analysis

A small microfluidic channel was made by sandwiching a cleaned quartz slide and a glass coverslip using double-side tape. Beforehand, two holes were drilled into the quartz slide and afterwards tubing was attached to them to act as inlet and outlet ports that were used for flowing buffer and introducing the RNA. The two RNA oligonucleotides containing the Cy3, Cy5 and biotin modifications were mixed at a concentration of 2 µM each in 50 mM HEPES, pH 7.2, 100 mM KCl (1x buffer). The RNA was heated at 90 °C for 2 min, left at RT for 30 s to cool down, followed by addition of 2 mM $MgCl_2$ and slow cooling down to RT over 15 min (Appendix Figure 6.3). This annealed RNA stock was used for carrying out smFRET experiments, for which a 15-25 pM RNA solution was made by diluting the 2 μM annealed stock using 1x buffer.

100 μL of diluted RNA solution was flowed onto the quartz slide coated with biotinylated-BSA and streptavidin for immobilization and incubated for 2 min. Any unbound RNA was washed off using 1x buffer. This step also washes off small (low nM) amounts of residual Mg^{2+} present in the stock. For experiments in the absence of divalents, 0.1 mM EDTA was included in the 1x buffer to chelate any contaminating divalent ions. Mg^{2+} titration experiments were performed on the same slide after washing off the EDTA using 1x buffer. We included an enzymatic oxygen scavenging system (OSS) containing 50 nM protocatechuate-3,4-dioxygenase and 5 mM protocatechuic acid in 1x buffer to prolong the longevity of the fluorophores. In addition, 2 mM Trolox (6-Hydroxy-2,5,7,8-tetramethylchromane-2-carboxylic acid) was included in the imaging buffer to suppress photoblinking of the dyes.

All smFRET movies were acquired at ~16 Hz, unless otherwise specified, using a prism-based total internal reflection fluorescence (TIRF) microscope with an intensified CCD (I-Pentamax, Princeton Instruments) or sCMOS camera (Hamamatsu ORCA-Flash4.0 V3), essentially as previously described[83,128]. Cy3 on the RNA was excited using a 532 nm laser and the emission from both Cy3 (donor) and Cy5 (acceptor) were simultaneously detected side-by-side on the camera. Towards the end of all movies, the Cy5 was directly excited using a 640 nm laser to check for the presence of the acceptor (Cy5) fluorophore. This helps in distinguishing the low-FRET (~0.1 FRET) states we observe from ~0 FRET states due to the absence or photobleaching of Cy5 in Cy3-labeled molecules. Raw movies were analyzed using IDL (Research Systems) to extract the time traces for all spots in Cy3 and Cy5 channels. Single-molecule traces were then visualized using MATLAB and only those with a minimum combined intensity (Cy3 + Cy5 intensity) of 300, showing single-step photobleaching of the dyes, a signal-to-noise ratio of >3, and longer than 6 s were selected for further analysis. Selected traces were then background-subtracted to correct for cross-talk and (minimal) bleed-through. We calculated the FRET ratio as $I_A/(I_A+I_D)$, where I_A and I_D are the background-corrected intensities of the acceptor (Cy5) and donor (Cy3), respectively. FRET histograms were made using the first 50 frames (3 s) of all traces (at least 200) in a given condition and fit with a sum of Gaussians using OriginPro 8.5. For kinetic analysis, traces were idealized with a two-state model corresponding to undocked (low-FRET) and docked (high-FRET) states using the segmental k-means algorithm in QuB software as previously described[114,129]. Cumulative dwell-time histograms were plotted from all extracted dwell times and fit with single- or double-exponential functions using OriginPro 8.5 to obtain the lifetimes in the undocked (τ_{undock}) and docked (τ_{dock}) states. Rate constants of docking and undocking were then calculated as $k_{dock} = 1/\tau_{undock}$ and $k_{undock} = 1/\tau_{dock}$. For the double-exponential fits, kinetics were calculated similarly using both the short and long dwell lifetimes to obtain the fast and slow rate constants, respectively. The idealized smFRET traces were used for creating transition occupancy density plots (TODPs), which show the fraction of traces/molecules that exhibit a given type of transition at least once[114]. In TODPs, dynamic traces showing a FRET transition (regardless of the number of transitions in that trace) and static traces (with no transitions over the entire trace) are weighted equally, avoiding over-representation of the traces with fast transitions. This feature is distinct from the traditional transition density plots (TDPs) that have been used to show all FRET transitions in a population as off-diagonal contours in the heat map. By contrast, TODPs are meant to show the fraction of different kinds of molecular behaviors and highlights slow or rare transitions in a population of traces. In addition, the improved TODPs, as shown here, also include static traces, which show up as on-diagonal contours that are often critical in understanding the molecular mechanism of riboswitches[130,131]. Thus, the on-diagonal contours in these TODPs (at ~0.2 and ~0.7 FRET values) do not represent any transitions but rather highlight static traces that lack any FRET transitions over their entire length and are counted exactly once in the TODP heatmap.

2.5.4 SiM-KARTS assay

The SiM-KARTS assay to probe P1.1 accessibility was performed on a prism-based TIRF microscope as described previously[65]. 20 pM of the riboswitches was bound to the slide and the Cy5 on the riboswitch was first photobleached by excitation with a red laser. Kinetic measurements of P1.1 accessibility was performed in the presence of 10 nM of a Cy5-labeled DNA oligonucleotide and under varying concentrations of the metal ions in the 1x buffer with OSS also used for smFRET. The 5’ Cy5-labeled DNA oligonucleotide is complementary to the 12-nt sequence (5’-GAGACCAGGGAU-3’) of the 3’-half of L1 and P1.1 in the smFRET RNA construct (Figure 2.2). Movies were acquired at 100 ms time resolution under dual excitation with both a green and red laser to localize the Cy3-labeled surface-bound riboswitch molecules and to image the short binding events of the Cy5-labeled oligonucleotide probe, respectively. The Cy5 versus intensity time traces were idealized to a two-state model using the segmental k-means algorithm implemented in QuB to obtain the bound and unbound dwell times. The unbound and bound lifetimes were obtained by single-exponential fits to the cumulative dwell time distributions, the inverse of which gives the rate constants k_{on} and k_{off}, respectively. The errors of the rate constants were estimated using a bootstrapping algorithm by taking three random subsets from all the traces in a condition and calculating the rate constants. The means and standard deviations from the three calculated rates are reported in Figure 2.14.

Chapter 3. The Conformational Dynamics of a Mn^{2+} Sensing Riboswitch are Modulated by the Paused RNA Polymerase[3]

3.1 Overview

Riboswitch regulation of gene expression happens by precise recognition of the cognate ligand alongside multiple other cellular factors. Folding of transcriptional riboswitches is kinetically controlled during transcription, where the structure of riboswitch is determined beyond ligand recognition and decided by events like transcriptional pausing involving RNA polymerase (RNAP) and transcription factors for efficient regulation of the associated genes. However, the riboswitch folding mechanism during transcriptional pausing and its influence over a paused RNAP is not well understood. Using a combination of single-molecule and biochemical techniques, we determined the mechanism of an Mn^{2+} ion sensing riboswitch from *L. lactis* folded co-transcriptionally in the presence of a paused RNAP. Using biochemical assays, we characterized the pause as a class I pause. Using single-molecule FRET and colocalization assays, we showed that the riboswitch folding is not only dependent on the binding of Mn^{2+} ions, but also on the presence of a RNAP, as well as on transcriptional factors, like NusA. We identified a previously unknown intermediate conformation of the riboswitch that occurred in the presence of a paused RNAP. While binding of Mn^{2+} ion to the riboswitch core was observed to destabilize the intermediate conformation, and thus promoting stable docking through P1.1 stabilization, the presence of NusA prompted this conformation's stabilization over the stabilization of P1.1 to promote transcription termination. Our results demonstrate that class I pausing in co-transcriptional folding of riboswitches acts as a decision point to determine the fate of transcription regulation by an associated riboswitch.

---

[3] Some contents of this chapter are adapted from a manuscript of the same title under preparation, by Shiba S. Dandpat, Adrien Chauvier, Rosa Romero and Nils G. Walter. S.S.D., A.C. and N.G.W. designed the experiments; S.S.D., and A.C. performed most of the experiments, analyzed the data and wrote the manuscript.

3.2 Introduction

Transcriptional pausing of RNA Polymerase (RNAP) is a ubiquitous phenomenon that acts as a control gateway in the gene regulatory process[52,54,132]. Within bacteria, where the transcription and translation machineries can couple, mechanisms like pausing of RNAP help the ribosome to catch up, allowing synchronization of transcription and translation[133]. Events such as transcriptional pausing during elongation also determine the fate of transcription elongation as well as termination[134]. Among these essential roles of transcriptional pausing, one of its key functions that determine the fate of RNA is its interplay with co-transcriptional folding of RNA structure during transcription and while emerging from the RNAP exit channel[23,135]. Sequence-specific transcriptional pausing has been identified for its role in promoting biologically active RNA structures by allowing time for the nascent RNA to fold[54,136,137]. More specifically, pauses regulating *trp* operon have been observed to allow time to form or rearrange nascent RNA structures[138]. Conversely, co-transcriptional folding of nascent RNA structure has also been identified to affect transcription speed and the fate of transcriptional pausing[57]. In addition, the presence of transcription factor proteins (like NusA, NusG, RfaH, etc.) play a key role in determining the lifetime of the pause, speed of RNAP, and ultimately the fate of transcription, hence influencing outcomes of co-transcriptional folding of nascent RNA in this process[136]. NusA is a well-known transcription factor that plays a key role in influencing pausing[55]. NusA is a multi-domain protein known to enhance hairpin stabilized pauses and promote transcription termination by interacting with the transcribing RNA as well as the RNAP exit channel[55,139,140]. Due to such complexities of the transcription machinery involving multiple components beyond just transcribing RNA, the interplay between the kinetics of transcription elongation and that of co-transcriptional folding as well as their respective roles in determining the conformational states of the transcribing RNA and, ultimately, gene regulation have rarely been addressed.

Riboswitches are a class of regulatory elements in bacterial messenger RNA (mRNA) that regulate gene expression by changing their structure in response to the rapidly changing environmental conditions[141,142]. Riboswitches exquisitely determine the outcome of transcription or translation in the expression platform by sensing cellular metabolites (ranging from ions to large molecules) acting as specific cognate ligands in their aptamer[97]. During transcriptional control, they modulate the formation or disruption of the terminator hairpin by adopting alternating structures co-transcriptionally[23]. Isolated RNA aptamers and their conformational dynamics under equilibrium conditions have been extensively studied[43,44,47]. However, due to the convoluted effect of riboswitch co-transcriptional folding, the kinetics of such folding in the presence of the transcription machinery like RNAP has been difficult to explore. Recent studies with riboswitches in the presence of an RNAP elongation complex (EC) have revealed that the conformational changes in the riboswitches are heavily influenced by such machinery[23,38,137], demonstrating the importance of such systems for the mechanistic understanding of co-transcriptional folding of RNA. Studies on the thiamine pyrophosphate (TPP) riboswitch using single-molecule FRET (smFRET) have shown that co-transcriptional folding of the riboswitch is kinetically controlled, where a brief window between ligand binding and transcriptional pausing determines the fate of downstream gene expression[57]. Additional studies on fluoride (F^-) and ZTP sensing riboswitches using co-transcriptional SHAPE-seq revealed the interplay of kinetic control between co-transcriptional folding and ligand binding that determines the fate of riboswitch regulated gene expression[38,143]. However, there is still a lack of kinetic and mechanistic insights into transcription elongation events like pausing and the role of additional factors like NusA in determining the fate of riboswitch folding.

The *yybp-ykoy* RNA motif is a widespread class of riboswitch that is found in bacteria, including many human and plant pathogens[106]. It is a highly structured RNA motif, identified to upregulate or downregulate a variety of genes at the level of transcription as well as translation, particularly those associated with Mn^{2+} ion homeostasis [110]. It does so by selectively sensing Mn^{2+} ions at a sub-millimolar level over predominantly available Mg^{2+} ions or any other divalent ions in the cell[82,106]. Crystal structure of the riboswitch from *Xanthomonas oryzae* (*X. oryzae*) and *Lactococcus lactis* (*L. lactis*) revealed two helices P1 and P3 stacked on top of each other, forming a four-way junction (4WJ) with two metal ion binding pockets formed by loops L1 and L3 with selective sensing Mg^{2+} and Mn^{2+} ions (Figure 3.1A)[82,106]. smFRET studies have revealed that while the presence of a physiological concentration of Mg^{2+} ion pre-organizes the riboswitch into a compact global structure, the capturing of Mn^{2+} ion stabilizes the "docked" conformation responsible for the up regulation via anti-terminator formation[82]. The docked conformation of the riboswitch aptamer is further stabilized by the formation of a stable P1.1 helix that otherwise switches dynamically between aptamer and terminator formation in the expression platform[82,106].

However, little is known about how the riboswitch folds during transcription. It is also unknown how the P1.1 switch helix is formed and stabilized to determine the outcome of transcription.

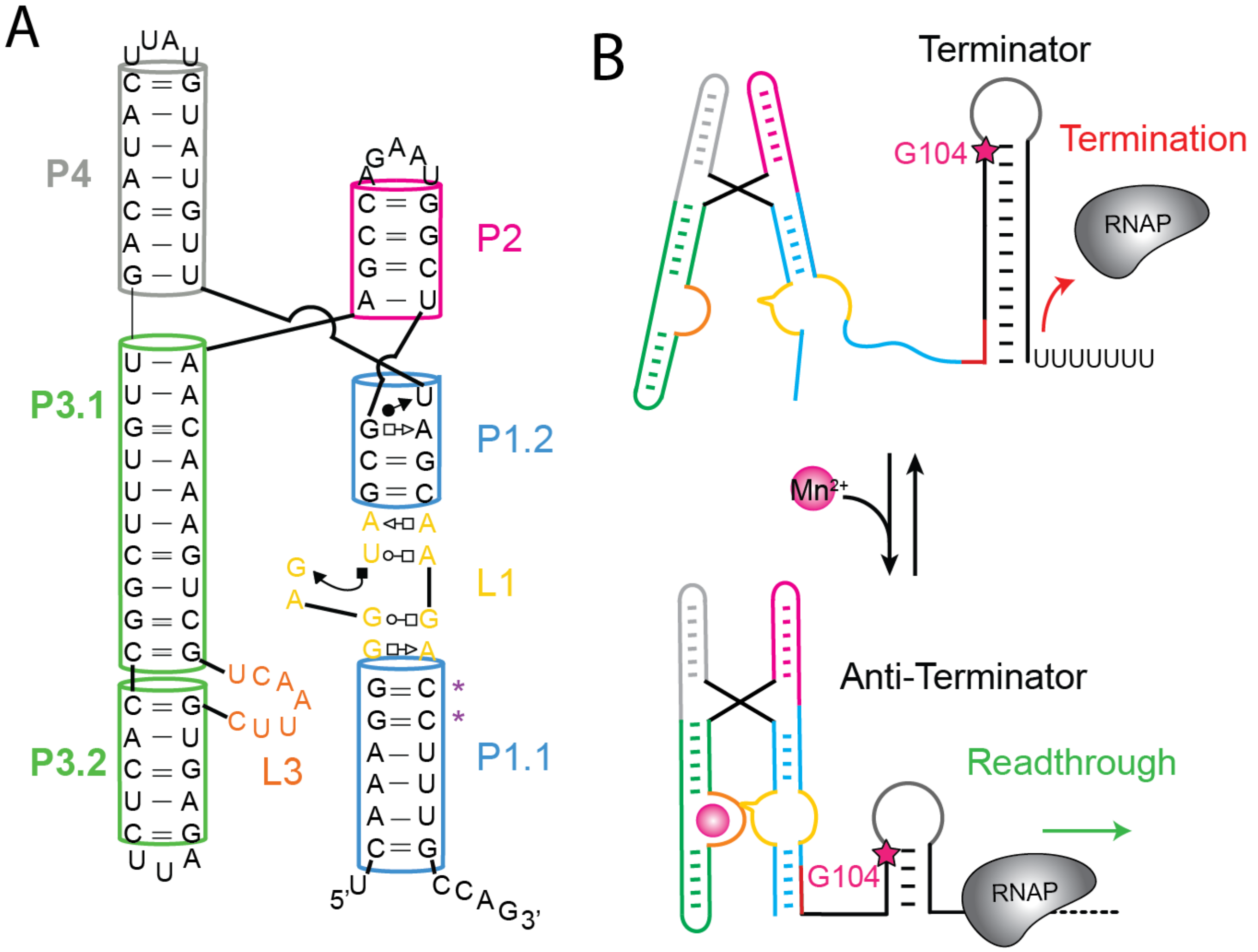


Figure 3.1 Schematic representation of the manganese (II) sensing riboswitch and its mechanism for transcription control.

(A) Secondary structure of the Mn^{2+} sensing riboswitch from *L. lactis* forms a 4WJ with two helices, P1 and P3 stacked upon each other. Ligand binding core L1 and L3 allows binding of Mg^{2+} and Mn^{2+} to their core respective. The riboswitch is extended beyond P1.1 until G104. The star on the right to P1.1 indicates the site for mutations to disrupt P1.1. (B) Schematic representation of transcription regulation by Mn^{2+} sensing riboswitch shows when Mn^{2+} is not bound to the ligand-binding core, the riboswitch adopts an open or undocked conformation that forms of terminator hairpin over P1.1 helix. With the addition of Mn^{2+} ion, the riboswitch adopts a closed or docked conformation by forming a stable P1.1 and disrupting the terminator to the anti-terminator hairpin, promoting transcription readthrough.

Current studies with the isolated Mn^{2+} sensing riboswitch have revealed the specificity of Mn^{2+} ion binding to the riboswitch is essential for the upregulation of the associated genes[82,124]. However, the influence of the riboswitch to control transcription outcomes happen *in-vivo* when the riboswitch is folding co-transcriptionally and the folding pathway is kinetically controlled by events of transcription elongation[144]. To understand the role events such as transcriptional pausing have on determining the conformational dynamics of the Mn^{2+} sensing riboswitch folded co-transcriptionally as well as under equilibrium conditions, we investigated the Mn^{2+} sensing riboswitch from *L. lactis* by observing it within a paused elongation complex (PEC). By utilizing a combination of single-molecule methods and biochemical assays, we investigated the role of RNAP and ligand Mn^{2+} ions in determining the riboswitch conformation and formation/destabilization of the P1.1 region and found that folding of the riboswitch occurs beyond the binding of Mn^{2+} and is significantly influenced by the paused RNAP. We further identified the role of a transcription factor, NusA, in determining the fate of riboswitch conformations in a PEC and its influence due to the binding of Mn^{2+} ions. Overall, our results provide kinetic and mechanistic insights into how co-transcriptional folding of a riboswitch and its structural changes occur beyond ligand binding and are influenced by factors of the transcription machinery including RNAP and transcription factor NusA, leading to a nuanced understanding of how a riboswitch in its native context regulates gene expression.

3.3 Results

3.3.1 Mn^{2+} binding to the aptamer stabilizes transcriptional pausing in the downstream expression platform

Structurally, Mn^{2+} sensing riboswitches resemble a 4-way junction (4WJ), with two helices P1 and P3, and loops L1 and L3 that undergo dynamic interactions to attain divalent metal ions responsive docking and undocking of the riboswitch (Figure 3.1A). In the absence of the Mn^{2+} ion, the riboswitch remains in an open or "undocked" conformation, where loop L1 and L3 are not occupied with divalent ions. This leads to the downregulation of gene expression promoted by the formation of a terminator hairpin, causing transcription termination[82,106]. In contrast, when Mn^{2+} ion is bound to L3, the riboswitch reorganizes its structure by stacking the loops L1 and L3 on top of each other through an A-minor interaction between L1, L3, and P3.1 and as a result stabilizing stem P1.1[82], which leads to disruption of the terminator to anti-terminator and allowing transcription readthrough (Figure 3.1B). From our earlier understanding of the mechanism of the Mn^{2+} sensing riboswitch and the potential role of RNAP in determining riboswitch structure[23,82,106], we hypothesized that transcriptional pausing of RNAP plays an important role in determining the riboswitch conformation and stabilization of P1.1 stem.

To test this hypothesis, we investigated to identify RNAP pause sites during the transcription of the riboswitch. We performed single-round *in vitro* transcription assays, where aliquots of the transcription reactions were taken at different time points (Figure 3.2A). We identified a long-lived pause site at position G104, downstream of the riboswitch aptamer. The pause site was observed to be located on the 5' arm of the terminator hairpin, where the RNAP paused with a half-life of ~194 s in the absence of ligand (Figure 3.2B). The addition of 0.5 mM Mn^{2+} ions enhanced the half-life of the pause by almost two-and-a-half times to ~523 s (Figure 3.2B). This indicated that the formation of a stable structure in the 5' region of the riboswitch and stabilization of partially formed P1.1 was due to the binding of Mn^{2+} ion that enhances the pause half-life. This dependence on Mn^{2+} ions for the increase in pause half-life further showed that the pause at G104 is dependent on the structure of the riboswitch aptamer and its influence with Mn^{2+} binding.

As seen from Figure 3.1A, the pause site G104 is about ten bases away from the starting base of the P1.1 stem. During transcription elongation, ~7-9 bases of transcribing RNA are known to pair with the DNA template strand in the transcription bubble[145]. This means that in the present case when RNAP is paused at G104, a part of P1.1 can be expected to be inside the RNAP. To test the role of the P1.1 formation and stabilization due to the binding of Mn^{2+} ions and their role in enhancing pausing half-life, we mutated the two cytidine residues at positions 95 and 96 to two guanosine residues to disrupt the P1.1 stem region (Figure 3.2B). With the destabilization of P1.1 through such mutations, we found that the half-life of the pause at G104 was of the same order in the absence and presence of Mn^{2+} ions, hence showing no dependence of pausing of RNAP on Mn^{2+} ions when P1.1 is disrupted (Figure 3.2B). These data showed that the RNAP pausing half-life at G104 is influenced by the structure formation in the aptamer region, which is determined by the binding of Mn^{2+} ions. These observations corroborate with previous studies that showed that transcriptional pausing at specific locations can coordinate structural changes in riboswitch to coordinate gene regulation[135,146].

To characterize the nature of the pause further, we tested the *in vitro* transcription assays in the presence of a transcription factor, NusA. For the WT riboswitch, in the absence of Mn^{2+} ions, we observed a four-time increase in the pause half-life to ~926 s for the RNAP paused at G104 (Figure 3.2B). Upon addition of Mn^{2+} ions, we did not observe a significant difference in pause half-life, showing that the presence of NusA does not influence the pause half-life when the P1.1 is partially formed in the presence of the Mn^{2+} ion. To test the effect of NusA where riboswitch structure is disrupted due to mutation of the two cytidines residues at positions 95 and 96 to two guanosine to disrupt the P1.1 stem region, we observed that the pausing half-life was further enhanced both in the absence and in the presence of Mn^{2+} ions (Figure 3.2B). The observations show that, when the riboswitch is unstructured or when its structure was disrupted, NusA significantly enhanced the pausing of RNAP. NusA is well known to interact with RNAP and nascent RNA to induce pausing and promote transcription termination by facilitating RNA hairpin folding[147]. During transcription elongation, NusA is known to enhance the pause half-life for a hairpin dependent class I pause[55,148]. From our observation, as the pause at G104 is stabilized by the formation of riboswitch aptamer akin to a hairpin and enhanced by NusA, we classified the pause as a class I pause.

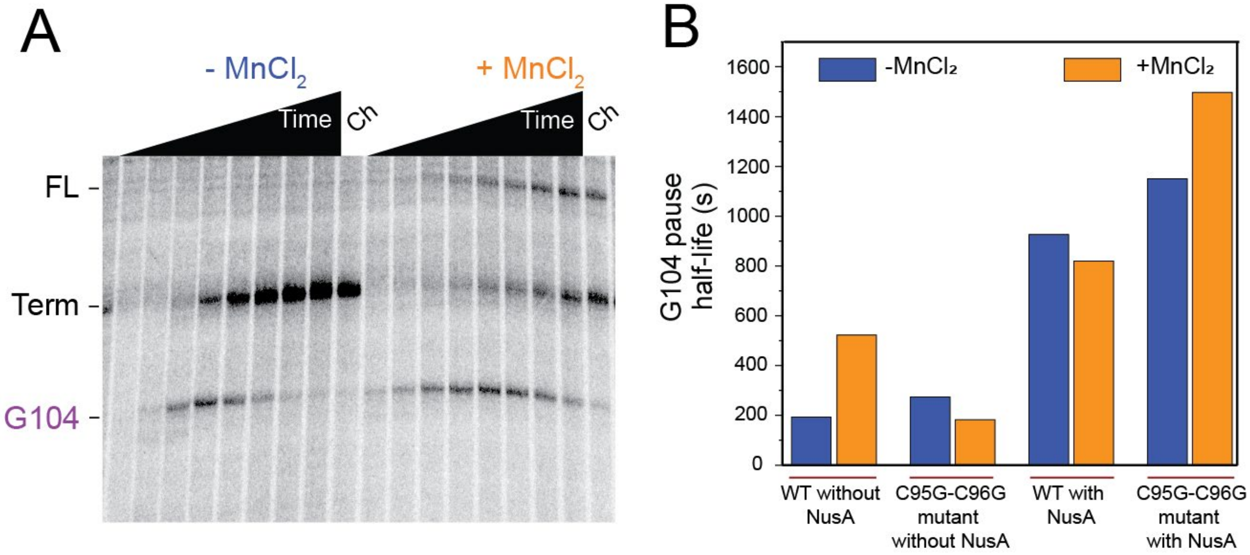


Figure 3.2 The Mn^{2+} sensing riboswitch regulate the outcome of transcription termination and pausing

(A) *In vitro* transcription assay of the Mn^{2+} sensing riboswitch from *L. lactis* in the presence and absence of $MnCl_2$ show that the RNAP is paused at G104. (B) Bar plots displaying the RNAP pause half-lives at G104 in the absence and presence of $MnCl_2$ for different co-transcriptionally folded riboswitch for the WT and C95G/C96G P1.1 mutant construct, as well as transcription, monitored in the absence and presence of transcription factor NusA.

3.3.2 Mn^{2+} binding stabilizes helix P1.1 during co-transcriptional folding of the riboswitch

After identifying the pause at G104 and the variation in its half-life due to the riboswitch structure, we next asked what role this pause plays in determining riboswitch structure, specifically at the P1.1 region. To answer this question, we prepared an *in vitro* transcribed riboswitch co-transcriptionally folded under native conditions wherein, upon reaching the pause suite, the RNAP was halted to form a PEC for single-molecule studies[137] (Figure 3.3A). We initiated transcription with a dinucleotide fluorescently labeled with a Cy3 fluorophore (AU-Cy3) and a subset of rNTPs to synchronize the transcription. Upon addition of all rNTPs, an active elongation complex stalled at pause site G104 is generated in the presence of a biotin-streptavidin roadblock at the end of the DNA template (Mn 104 PEC) (Figure 3.3A, Methods). The stalled PEC was immobilized to a biotin-PEG and streptavidin-coated quartz surface for imaging via prism-TIRF microscopy (Figure 3.3B). A buffer solution with Cy5 labeled short DNA oligonucleotide with 8 base pair complementarity with the 3' end of the riboswitch involving P1.1 region was added to monitor the kinetics of its formation and disruption under the influence of Mn^{2+} binding to the riboswitch in the PEC via Single-Molecule Kinetic Analysis of RNA Transient Structures (SiM-KARTS) assay[20,149]. With simultaneous excitation of Cy3 and Cy5 fluorophores, we monitored the continuous binding and dissociation of the Cy5-oligo to the immobilized RNA detected as spikes of fluorescence intensity in the red channel, colocalized with the green channel (Figure 3.3C & D). Two-state idealization of the binding and unbinding events for each molecule resulted in cumulative bound ($t_{unbound}$) and unbound (t_{bound}) times, which were used to determine the corresponding association (k_{on}) and dissociation (k_{off}) rate constants of the SiM-KARTS probe to its target (Figure 3.3E & F, Methods).

In the absence of Mn^{2+} ions, the SiM-KARTS probe binds and dissociates from the P1.1 region of the target RNA with two components for both the association (k_{on}) and dissociation (k_{off}) rate constants, resulting in a fast (k_{fast}) and slow (k_{slow}) component of each rate constants. The association rate constants were observed to be almost equal in contribution, resulting in $k_{on,\,fast}$ of $4.63 \times 10^6\ M^{-1}s^{-1}$ (49%) and $k_{on,slow}$ of $0.65 \times 10^6\ M^{-1}s^{-1}$ (51%), and the dissociation constants resulting in $k_{off,fast}$ of $1.31\ s^{-1}$ (44%) and $k_{off,slow}$ of $0.13\ s^{-1}$ (56%) (Figure 3.3E & F).

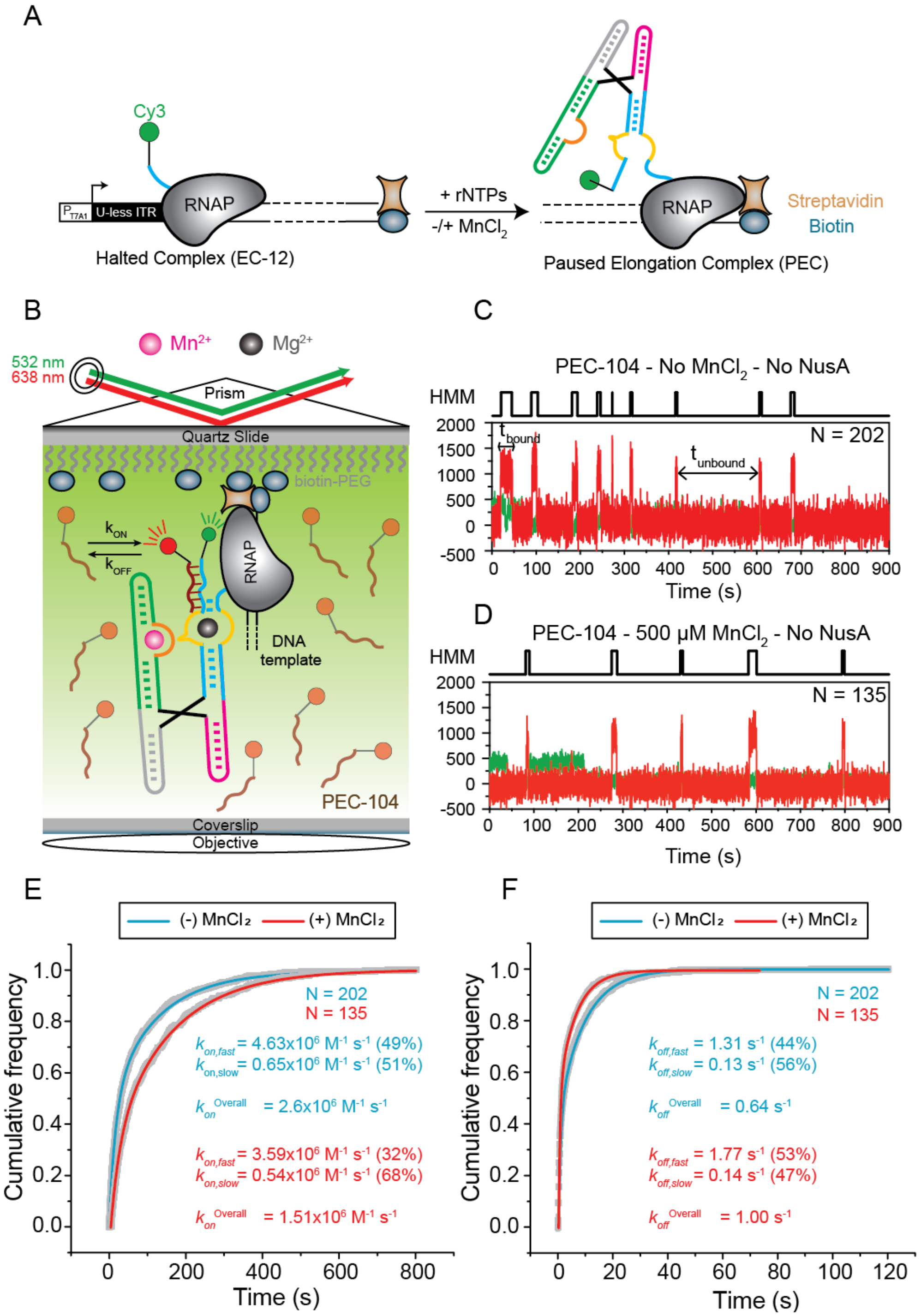


Figure 3.3 Schematic representation of SiM-KARTS assay revealing the formation and stabilization of P1.1 region within the PEC

(A) Fluorescently labeled Paused Elongation Complexes (PECs) are prepared through in vitro transcription using *E. coli* RNAP. The halted complex at G104 (Mn 104 PEC) is prepared through the addition of a dinucleotide labeled with Cy3 (AU-Cy3) and UTP deprivation (ATP/CTP/GTP). The biotin-streptavidin interaction on the 3'-end of the DNA template constitutes a stable transcriptional roadblock to stall the RNAP at the desired position and for subsequent immobilization on the microscope slide. (B) SiM-KARTS assay for studying colocalization of Cy5-labeled DNA probe (SiM-KARTS) with the co-transcriptionally folded RNA transcribed until the pause at 104 (Mn 104 PEC). Repeated binding and dissociation of the SiM-KARTS probe targeting the P1.1 switch helix were monitored through direct excitation of Cy3 and Cy5 fluorescent dyes. (C-D) Representative time traces for a single molecule colocalized with Cy3 and Cy5 signal showing binding and dissociation of the Cy5 labeled SiM-KARTS probe (in red) monitored in the absence (C) and presence (D) of 0.5 mM $MnCl_2$. Two-state idealization was performed for the individual trace to determine the probe bound (t_{bound}) and unbound ($t_{unbound}$) time to the riboswitch are shown in black at the top of each trace. Colocalized green signal (Cy3-RNA) was used as a reference to determine the presence of a single PEC at each localized position. N represents the total number of molecules identified and is further used to determine rate constants. (E-F) Plots displaying cumulative unbound (E) and bound (F) dwell times of the SiM-KARTS probe binding in the absence (blue) and the presence (red) of 0.5 mM $MnCl_2$. Bi-exponential association ($k_{on,fast}$ and $k_{on,slow}$) and dissociation ($k_{off,fast}$ and $k_{off,slow}$) rate constants alongside their normalized contributions for the probe binding are represented for each condition. Weighted average rate constants ($k^{overall}$) are shown for each condition.

The two binding and dissociation behavior of the SiM-KARTS probe suggests two types of alternate structures adopted with almost equal probability. The alternate structure of the P1.1 region within the PEC may be due to the fully disrupted or partially formed P1.1 through the available GG-CC interaction at top of the P1.1 region, resulting in differential accessibility for the probe to bind. While the faster association rate $k_{on,fast}$ of 4.63×10^6 $M^{-1}s^{-1}$ (51%) can be attributed to probe binding to the strand in the disrupted P1.1 region, the slower association rate, $k_{on,slow}$ of 0.65×10^6 $M^{-1}s^{-1}$ (49%) can be attributed to the probe binding to the partially formed P1.1. As the addition of Mn^{2+} ion to the riboswitch is known to stabilize the P1.1 stem[82], upon the addition of 0.5 mM $MnCl_2$ during co-transcriptional folding of the riboswitch influenced both the association and dissociation rates (Figure 3.3E & F). Under such conditions, the association rate resulted in ~23% decrease in $k_{on,fast}$ to 3.59×10^6 $M^{-1}s^{-1}$ (32%) and ~17% decrease in $k_{on,slow}$ to 0.54×10^6 $M^{-1}s^{-1}$ (51%). The weighted average association rates ($k_{on}^{overall}$) reduced by ~42% from 2.6×10^6 $M^{-1}s^{-1}$ to 1.5×10^6 $M^{-1}s^{-1}$ with the addition of Mn^{2+} ions to the riboswitch, further supporting the role of Mn^{2+} on the stabilization of P1.1. It is also important to note that while the changes in constant rate values are small, the contribution of slower association rate constant component increased from 51% to 68%, showing that the riboswitch in the presence of Mn^{2+} ion can stabilize the P1.1

region with only the top part of the stem available for base-pair formation. The dissociation rate constants with the addition of Mn^{2+} resulted in $k_{off,fast}$ of 1.77 s^{-1} (44%) and $k_{off,slow}$ of 0.14 s^{-1} (56%). A ~35% increase in the weighted average dissociation rate constants ($k_{on}^{overall}$) from 0.64 s^{-1} to 1.0 s^{-1} further shows that the addition of the Mn^{2+} ions to the riboswitch not only disrupt the SiM-KARTS probe binding by stabilizing partially available P1.1 but further destabilizes the bound SiM-KARTS probes to the P1.1 region.

To verify the effect of P1.1 stabilization on the SiM-KARTS probe binding, we prepared an *in vitro* transcription assay as earlier with the transcribed riboswitch mutated the 2Cs at position 95 and 96 to 2Gs to disrupt the formation of any P1.1 and allowed the SiM-KARTS probe to bind and dissociate to this construct (Figure 3.4A, B&C). For the mutant riboswitch embedded in a PEC halted at G104 (Mn 104 PEC MP1.1) and co-transcriptionally folded in the absence of Mn^{2+}, the association constants resulting in $k_{on,fast}$ of 8.04 x 10^6 $M^{-1}s^{-1}$ (13%) and $k_{on,slow}$ of 0.41 x 10^6 $M^{-1}s^{-1}$ (87%), and the dissociation constants were resulting in $k_{off,fast}$ of 0.69 s^{-1} (87%) and $k_{off,slow}$ of 0.05 s^{-1} (13%). For the mutant riboswitch (Mn 104 PEC MP1.1) co-transcriptionally folded in the presence of Mn^{2+}, the association constants resulting in $k_{on,fast}$ of 8.14 x 10^6 $M^{-1}s^{-1}$ (10%) and $k_{on,slow}$ of 0.38 x 10^6 $M^{-1}s^{-1}$ (90%),and the dissociation constants resulting in $k_{off,fast}$ of 0.71 s^{-1} (89%) and $k_{off,slow}$ of 0.05 s^{-1} (11%). This shows that unlike the WT, mutation of the P1.1 at 95 and 96 positions did not influence the binding kinetics of the SiM-KARTS probe to P1.1 in the absence or presence of Mn^{2+}. This may possibly be because the mutation of P1 could destabilize the binding of Mn^{2+} to the riboswitch[82].

To understand the role of RNAP in the ligand-dependent modulation of the SiM-KARTS probe binding to the partially available P1.1, we transcribed an RNA template until position 96 and released it from the elongation complex in the absence of a streptavidin-biotin roadblock as described earlier. This released RNA transcript (Mn 96 RT) was captured using a biotinylated DNA capture probe complementary to the 3' region of the riboswitch for attachment to the slide surface. Using the SiM-KARTS assay, we further probed the accessibility of the partially formed P1.1 region until position 96, similar to the Mn 104 PEC construct (Figure 3.4A, D & E). In the absence of Mn^{2+}, the association rate constants resulted in $k_{on,fast}$ of 9.96 x 10^6 $M^{-1}s^{-1}$ (40%) and $k_{on,slow}$ of 0.36 x 10^6 $M^{-1}s^{-1}$ (60%), and the dissociation constants resulted in $k_{off,fast}$ of 1.08 s^{-1} (69%) and $k_{off,slow}$ of 0.07 s^{-1} (31%).

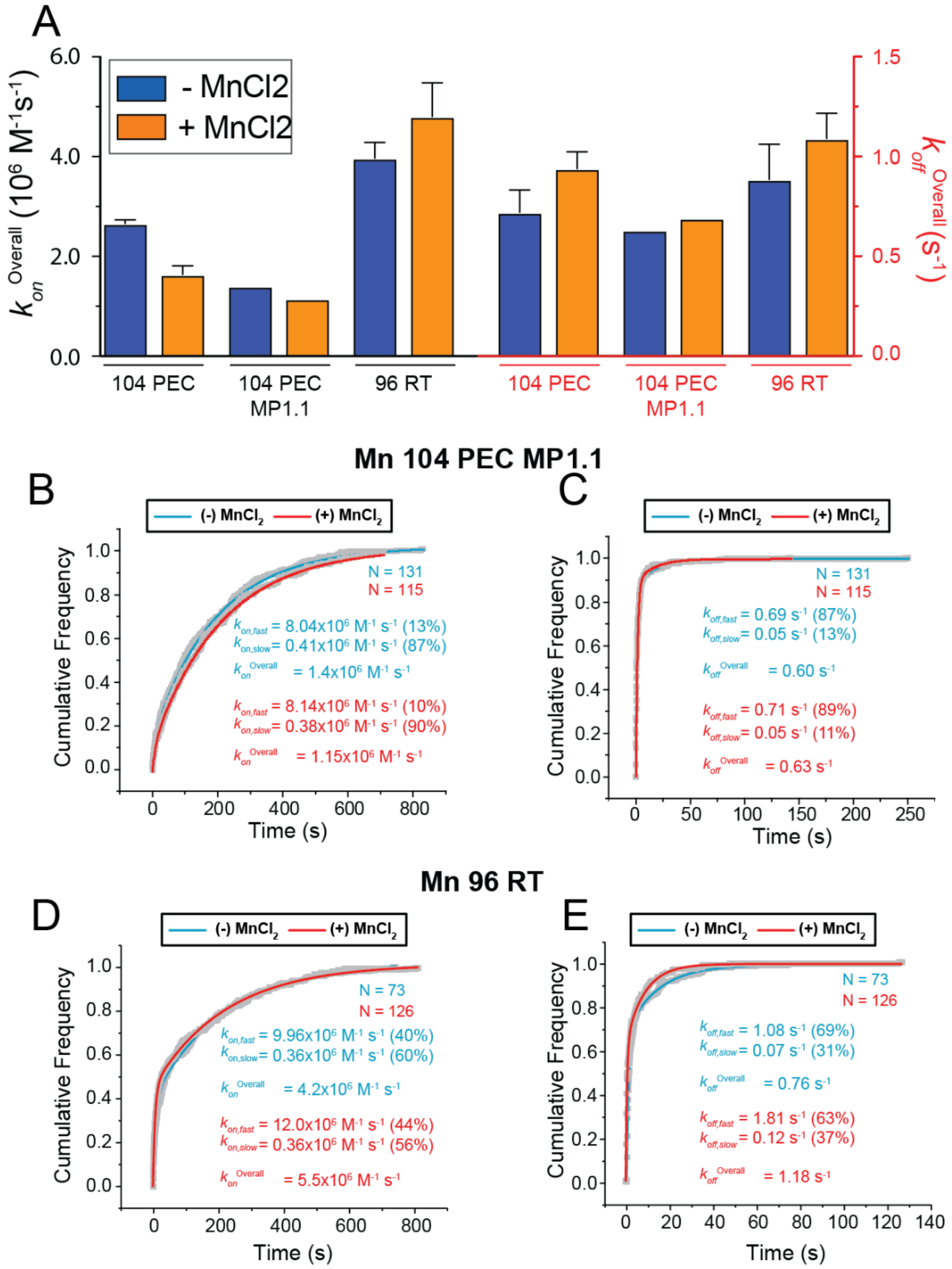


Figure 3.4 SiM-KARTS probe binding showed that riboswitch folding depends on the presence of Mn^{2+} as well as RNAP

(A) Compaction of overall association (left) and dissociation (right) rate constants three types of construct co-transcriptionally folded without and with Mn^{2+} ions. (B, D) Cumulative frequency plots showing association (k_{on}) rate constants for riboswitch transcribed until C96 and released from RNAP for SiM-KARTS assay on the riboswitch transcribed until G104 with C95G/C96G mutation in P1.1 and halted in the PEC (Mn 104 PEC MP1.1), and riboswitch release transcript (Mn 96 RT). (C, E) Cumulative frequency plots showing dissociation (k_{off}) rate constants for riboswitch transcribed until C96 and released from RNAP for SiM-KARTS assay on the riboswitch transcribed until G104 with C95G/C96G mutation in P1.1 and halted in the PEC (Mn 104 PEC MP1.1), and riboswitch release transcript (Mn 96 RT).

Compared to the probe binding to the Mn 104 PEC, the Mn 96 RT folded in the absence and presence of Mn^{2+} ions resulted in a faster association rate ($k_{on,fast}$) as well as the overall association rate ($k_{on}^{overall}$) increased by almost two-fold (Figure 3.4A). This showed that the SiM-KARTS probe was more accessible to binding to the P1.1 region when RNAP was absent. From the weighted average dissociation rates ($k_{off}^{overall}$), the absence of RNAP slightly increased the dissociation rates. However, their influence with respect to Mn^{2+} remained very similar.

Taken together, our SiM-KARTS data with co-transcriptionally folded riboswitch showed that the RNAP paused at position 104 plays a major role in the stabilization of a partially formed P1.1. While the addition of Mn^{2+} promotes further stabilization of P1.1, its influence was overshadowed by the RNAP hosting the transcribing riboswitch at its exit channel. In the overall determination of riboswitch structure and P1.1 formation was observed to be influenced by the presence of Mn^{2+} ion as well as RNAP.

3.3.3 smFRET histograms revealed the distribution of conformational states of the riboswitch present within a PEC

Previous studies of isolated Mn^{2+} sensing riboswitch aptamers have been observed to sample two predominant conformational states, where the low FRET represented an undocked state and the high FRET state represented a docked state[49,82]. To study the conformational dynamics of the riboswitch within in PEC, we used smFRET by labeling the two helices of the riboswitch and attaching it to a paused EC at G104 (Figure 3.5). To immobilize the riboswitch to the microscope slide surface, we prepared the PEC by annealing the two strands of the riboswitch with an artificial DNA bubble and added it to biotinylated RNAP[23]. The entire complex was prepared in such a way that only PEC containing fluorophore-labeled RNA attached to DNA would be observed in the field of view (Figure 3.5, Methods).

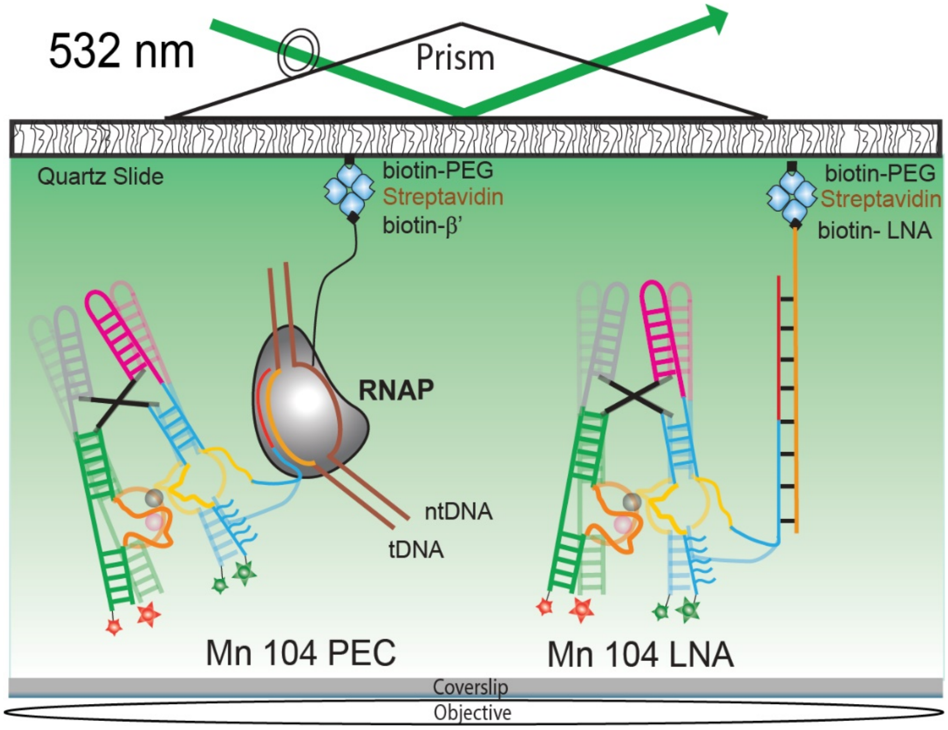


Figure 3.5 smFRET assay revealed the conformational dynamics of the Mn^{2+} sensing riboswitch immobilized through a PEC and LNA

Schematic representation of the smFRET assay using pTIRFM. In the Mn 104 PEC complex, the riboswitch was immobilized at pause site G104 and attached to a DNA bubble made by the template and non-template DNA. The whole complex was then added to a biotinylated *Eco* RNAP to reconstitute a complete paused elongation complex (Mn 104 PEC). In the Mn 104 LNA complex, the riboswitch was immobilized to the surface using a biotinylated LNA complementary to the 3' end of the riboswitch in a similar way as template DNA (Mn 104 LNA).

To monitor the global conformational dynamics of the riboswitch paused at G104 in a PEC (Mn 104 PEC), we used smFRET to monitor the fluorophores at the legs to helices P1 and P3. (Figure 3.5, Left). smFRET traces and their cumulative histograms monitored in presence of 100 mM KCl only showed that a slow sampling of two previously observed predominant FRET states at ~0.2 and ~0.6, with ~65% of the population in the low FRET state and ~35% population in the high FRET state (Figure 3.6A & B, Top panel). However, the riboswitch was also observed to occasionally sample an intermediate FRET state of ~0.45 stabilized over a few seconds (Figure 3.6A, Top trace). At a near-physiological concentration of Mg^{2+}, the riboswitch started to sample a more dynamic high FRET population (Figure 3.6A & B, Middle panel). The FRET histogram at 1 mM Mg^{2+} showed nearly equal populations in the high FRET (~0.7) and the low FRET regime

(~0.2), indicating that binding of Mg^{2+} to the riboswitch promotes more docked conformation for the riboswitch within the PEC. With the addition of 0.5 mM Mn^{2+} ions, the riboswitch sampled a more static high FRET population than in the absence of Mn^{2+}, representing more docked conformation (Figure 3.6A & B, Bottom panel). The FRET histogram showed a high FRET population (~0.7), which was further increased to ~69% from ~53% in the absence of Mn^{2+}, as compared to the low FRET population (~0.2), illustrating that addition of Mn^{2+} promotes more docked conformations over just Mg^{2+}, corroborating with earlier studies[82,124].

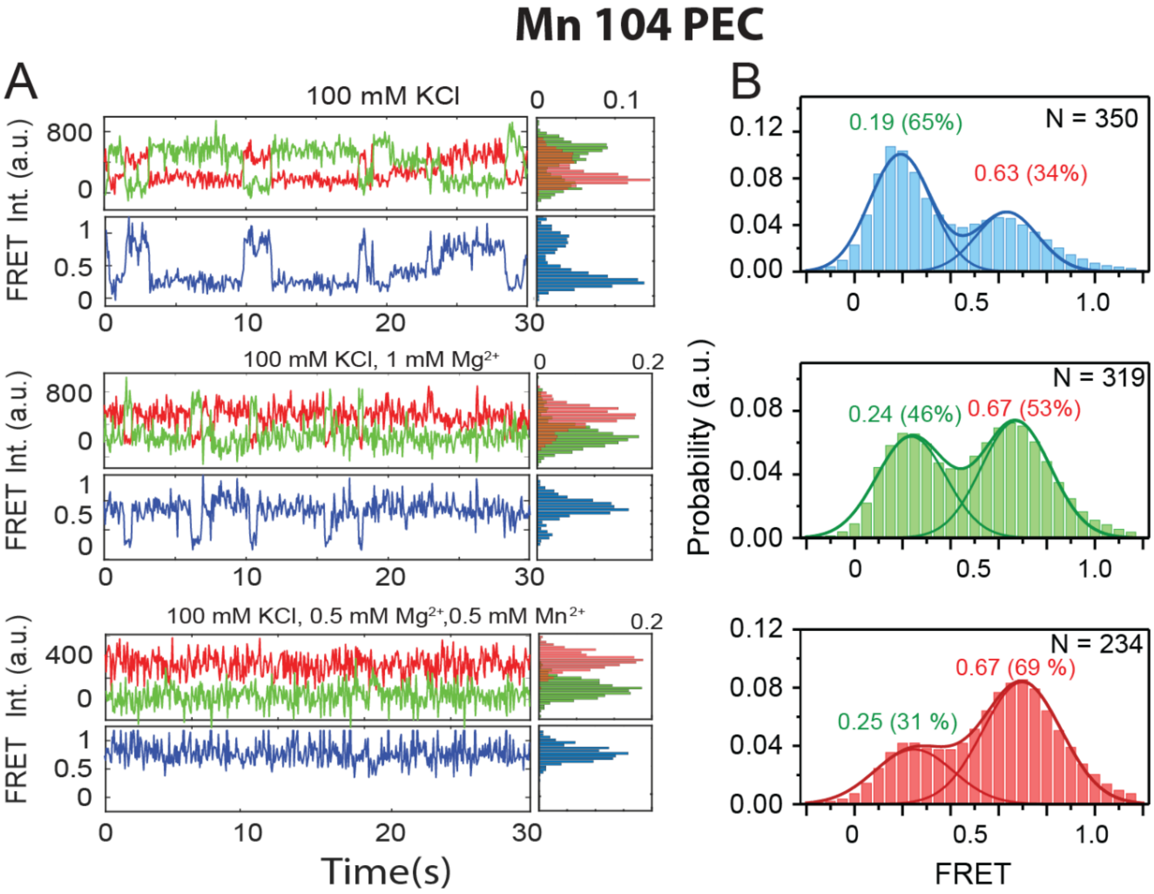


Figure 3.6 (A) Representative smFRET traces and (B) cumulative histograms for the Mn 104 PEC at different concentration of divalent ions (top to bottom) in the buffer: (i) no divalent ions, (ii) 1 mM Mg^{2+} only, and (iii) 0.5 mM Mg^{2+} and 0.5 mM Mn^{2+} respectively.

Green: Cy3, Red: Cy5, Black: FRET. The occupancy of the signals over time in each represented trace was shown in histograms to the right side of the trace. The histograms for cumulative FRET distribution were fitted with 2 Gaussian peaks to represent their predominant FRET population densities.

To assess the influence of the RNAP on the docking and undocking of the riboswitch, we used the same riboswitch aptamer extended up to the pause site G104 annealed to a biotinylated LNA capture strand. We ensured that the construct (Mn 104 LNA) mimicked the exact 8 base pair complementarity of the riboswitch with LNA, as it would have with the template DNA in the transcription bubble used in PEC experiments (Figure 3.5, Right). The smFRET traces monitored in presence of 100 mM KCl only showed less sampling of the high FRET population (Figure 3.7, Top panel). The FRET histogram showed more population in the low FRET regime (~0.2 with 76%) than the high FRET regime (~0.69 with 23%), indicating that the presence of RNAP induced more docking of the riboswitch. In the presence of 1 mM Mg^{2+}, the traces became more dynamic sampling high FRET populations, where a FRET histogram showed an increased high FRET population (~42%) than without any divalent ions (Figure 3.7, Middle panel). The addition of Mn^{2+} ions further promoted the population towards the high FRET (Figure 3.7, Bottom panel). The FRET histogram showed an increase in the high FRET population (~67%) than with only Mg^{2+} ions. It is interesting to note that when the FRET changes for Mn 104 LNA (Figure 3.7) were compared with Mn 104 PEC (Figure 3.6), for each condition, we observed that the presence of RNAP in a PEC has a significant influence on promoting higher FRET population even when the P1.1 stem was not completely formed with the transcription elongation paused at G104. The presence of RNAP in Mn 104 PEC increased the higher FRET population (~0.67) by ~10% compared to its absence in Mn 104 LNA, even in the absence of Mn^{2+} ions (Figure 3.6 & 3.7, Top and Middle panels). As the addition of Mn^{2+} is known to promote a high FRET population of the riboswitch, its addition in both cases promoted the stable docked conformation. Under such saturating condition of 0.5 mM Mn^{2+}, the observations in both cases were very similar, where the riboswitch was observed predominantly in the high FRET conformation (Figure 3.6 & 3.7, Bottom panels).

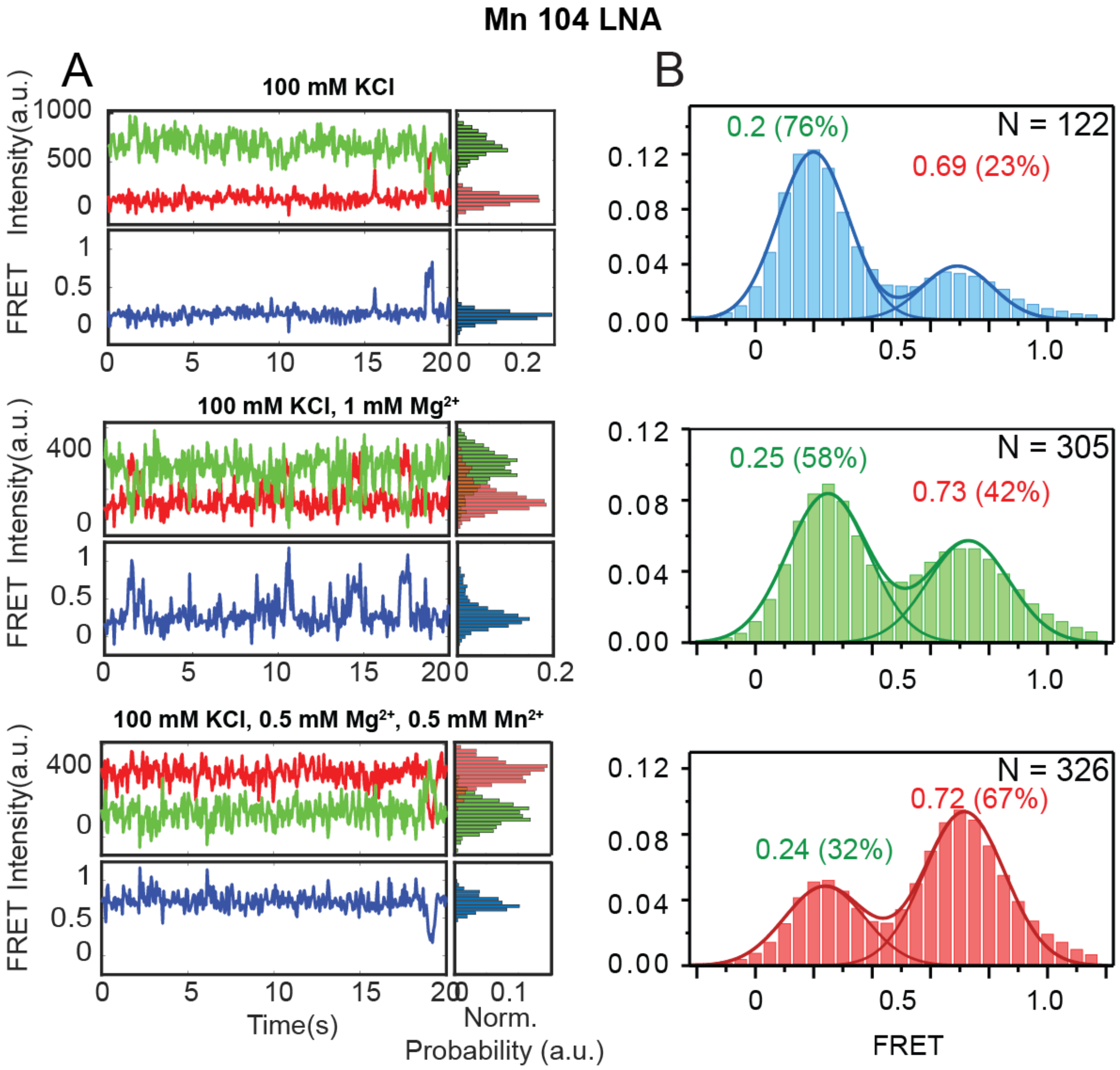


Figure 3.7 (A) Representative smFRET traces and (B) cumulative histograms for the Mn 104 LNA at different concentration of divalent ions (top to bottom) in the buffer: (i) no divalent ions, (ii) 1 mM Mg^{2+} only, and 0.5 mM Mg^{2+} and (iii) 0.5 mM Mn^{2+} respectively.

Green: Cy3, Red: Cy5, Black: FRET. The occupancy of the signals over time in each represented trace was shown in histograms to the right side of the trace. The histograms for cumulative FRET distribution were fitted with 2 Gaussian peaks to represent their predominant FRET population densities.

To further assess the role of P1.1 and how its formation and stabilization influence conformational dynamics of the riboswitch, we designed a construct that goes 10 bases beyond the pause (Mn 114 LNA) to resemble the condition where riboswitch is completely transcribed with fully formed P1.1. We immobilized the riboswitch (Mn 114 PEC) with a complementary LNA capture strand to attach it to the surface and recorded smFRET over time (Figure 3.8A). We hypothesized that under such a condition, the riboswitch would have moved outside of the exit channel of the RNAP and was away from the influence of RNAP. Under such conditions, in the absence of both Mg^{2+} and Mn^{2+}, we observed a predominantly low FRET population (~0.18) at 100 mM KCl (Figure 3.8, Top panel). With the addition of 1 mM Mg^{2+}, the riboswitch started sampling a higher FRET population. However, the majority of the population (~75%) were still in the low FRET (~0.22) regime (Figure 3.8, Middle panel), showing that Mg^{2+} alone is not able to promote the stable docked conformation when riboswitch aptamer is fully formed. With the addition of a saturating concentration of 0.5 mM Mn^{2+} along with 0.5 mM Mg^{2+} ions, the riboswitch promoted a predominantly high FRET population (~67%) (Figure 3.8, Bottom panel). These observations resembled earlier studies with isolated riboswitch aptamer[82,124] and showed that when the riboswitch was formed with fully available P1.1 and away from the influence of the RNAP, the riboswitch behaved as an isolated construct. Hence pausing of RNAP alongside the availability of P1.1 plays a key role in determining the conformational dynamics of the Mn^{2+} sensing riboswitch to precisely to determine the outcome of transcription.

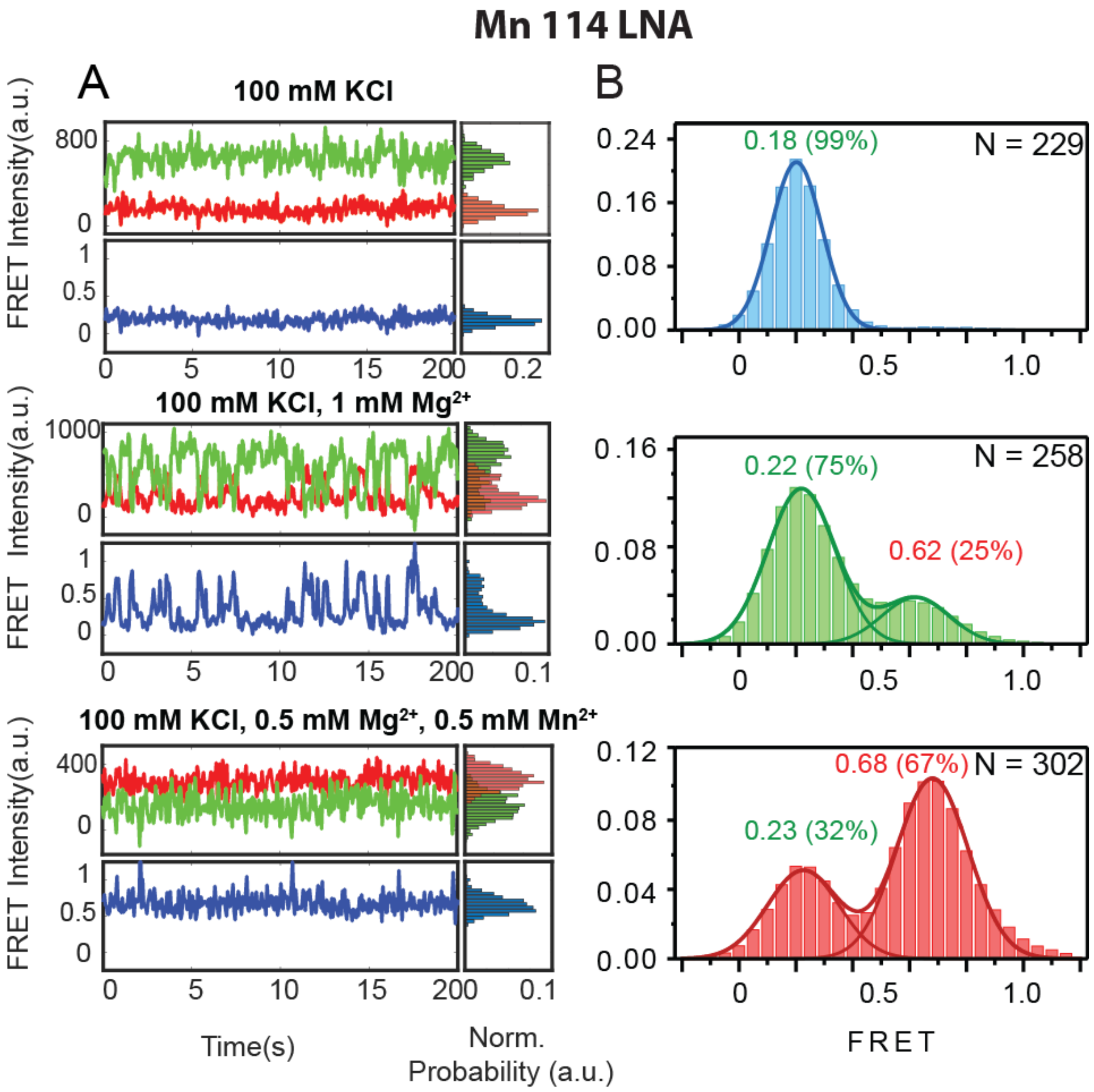


Figure 3.8 (A) Representative smFRET traces and (B) cumulative histograms for the Mn 114 LNA at different concentration of divalent ions (top to bottom) in the buffer: (i) no divalent ions, (ii) 1 mM Mg^{2+} only, and 0.5 mM Mg^{2+} and (iii) 0.5 mM Mn^{2+} respectively.

Green: Cy3, Red: Cy5, Black: FRET. The occupancy of the signals over time in each represented trace was shown in histograms to the right side of the trace. The histograms for cumulative FRET distribution were fitted with 2 Gaussian peaks to represent their predominant FRET population densities.

To compare the effect of PEC on the stable docked conformation, which is known to result from the formation of P1.1 required for preventing transcription termination[82], we calculated the total number of stable docked traces for each condition within the three constructs (Table 3-1). We found that Mn 104 PEC and Mn 104 LNA can promote a higher number of stable docked populations in the absence of Mn^{2+}, even with the partially formed P1.1 with only two CC:GG interactions at positions 95 and 96. In contrast, for Mn 114 LNA the riboswitch being away from any influence of the pausing and having a fully available P1.1 requires stabilization to achieve a docked conformation, which is reached only after Mn^{2+} ion is bound to the riboswitch. These observations of a higher population of stable docked conformation at higher Mn^{2+} concentration shows that the riboswitch can adopt a docked structure with stable P1.1 that can disrupt the terminator hairpin and allows transcription elongation to go forward, similar to a conformational selection pathway[43,124]. The presence of a higher fraction of the docked population in a PEC (Figure 3.6) as compared to the other complexes is possibly due to the steric constraints enforced by the RNAP machinery when the P1.1 is still present inside the exit channel while being transcribed. Such a sterically constrained environment of RNAP may be able to promote the folding of the riboswitch even with only a partially formed P1.1. Taken these data together with our earlier observations showed that, during transcriptional pausing at G104, the presence of Mn^{2+} ions can promote a docked conformation of the riboswitch to signal the transcription elongation forward.

Table 3.1 Percentage of stable-docked traces observed from smFRET at different divalent ion concentrations for (A) Mn PEC 104, (B) Mn 104 LNA, and (C) Mn 114 LNA.

Conditions	static docked traces	Total traces	Percentage static docked (%)
Mn PEC 104			
100 mM KCl Only	*47*	*350*	*13*
100 mM KCl with 1 mM Mg^{2+}	*63*	*319*	*19.7*
100 mM KCl with 0.5 mM Mg^{2+} & 0.5 mM Mn^{2+}	*84*	*234*	*35.9*
Mn LNA 104			
100 mM KCl Only	*14*	*122*	*11.5*
100 mM KCl with 1 mM Mg^{2+}	*55*	*305*	*18*
100 mM KCl with 0.5 mM Mg^{2+} & 0.5 mM Mn^{2+}	*149*	*326*	*45.7*

Mn LNA 114			
100 mM KCl Only	*2*	*229*	*0.9*
100 mM KCl with 1 mM Mg^{2+}	20	258	7.8
100 mM KCl with 0.5 mM Mg^{2+} & 0.5 mM Mn^{2+}	115	302	38.1

3.3.4 smFRET kinetics revealed transition between the states was accessed through an intermediate conformation for the riboswitch present within a PEC

To understand the kinetics of switching between riboswitch conformational states within the PEC at different concentrations of divalent ions, we performed three-state HMM modeling based on the three FRET states (L = Low (~0.2), I = Intermediate (~0.45), H = High (~0.75)) observed during the trace selection. We estimated the docking (k_{dock}) and undocking (k_{undock}) rate constants for six dwell times as well as their relative probability resulting from transitions between the three states. We represented the non-dynamic nature of the transitions as on-diagonal contours centered at ~0.2 (L) and ~0.75 (H), and the dynamic nature of the transitions as off-diagonal contours centered at ~0.2 (L), ~0.45 (I), and ~0.75 (H) in the transition occupancy density plot (TODP); such a TODP displays as a heat map the fraction of single-molecule traces that exhibit any given specific initial-to-final FRET transition at least once[150] (Methods). For the off-diagonal dynamic transitions, we estimated the k_{dock} and k_{undock} from single-exponential kinetic fits and their relative probability over the entirety of all dynamic transitions (Methods).

To determine the transition kinetics for the Mn 104 PEC, we calculated its on- and off-diagonal contours at different divalent ions concentration. At 100 mM KCl (Figure 3.9, Top panel), the riboswitch within PEC had the majority of the population (~40%) in the stable undocked conformation at ~0.2 FRET state, and the rest (~45%) dynamically sampled between the three conformational states. The transition between L and I states resulted in a slow k_{dock} of 1.5 s^{-1} and fast k_{undock} of 3.2 s^{-1}, as can be expected for the riboswitch to adopt an undocked conformation in the absence of any divalent ions. The transition between I and H states resulted in a faster k_{dock} of ~3.0 s^{-1} and slower k_{undock} of 2.2 s^{-1}, showing that the I state is very transient for the riboswitch. The transition between the predominant L and H resulted in a slower k_{dock} of ~2.2 s^{-1} and a much faster k_{undock} of ~3.5 s^{-1}. While the transition probability between the three states were almost similar, the transition from L was predominantly lower than I and H. The k_{dock} values also showed that while transitioning to H state, the riboswitch sampled transiently through the state I. The k_{undock}

values showed that the H state, while briefly accessible, was not very stable in the absence of divalent ions and often sampled through the I state. In the presence of 1 mM Mg^{2+} (Figure 3.9, Middle panel), the riboswitch within the PEC started to sample a more dynamic population (~70%) over stable undocked (~9%) and docked population (~20%). The transition between L and I states resulted in a slow k_{dock} of ~1.6 s^{-1} and a fast k_{undock} of ~2.8 s^{-1}, showing that the addition of Mg^{2+} didn't significantly impact the ability of the riboswitch to adapt L state over I state. The transition between I and H states resulted in a faster k_{dock} of ~3.1 s^{-1} and a slower k_{undock} of 1.7 s^{-1}, showing that the I state was transiently achieved by the riboswitch. The addition of Mg^{2+}, however, resulted in a slower transition from H to I. The transition between the predominant L and H resulted in a k_{dock} and k_{undock} values around ~2 s^{-1}. While the addition of Mg^{2+} did not affect the k_{dock} between L and H, it lowered the k_{undock}, showing that Mg^{2+} binding to the riboswitch while increased the dynamic transition between the states, it preferably stabilized the H state conformation more. Further addition of Mn^{2+} with Mg^{2+} ions (Figure 3.9, Bottom panel) increased the stable docked population (~35%) as compared to only Mg^{2+}, as well as lowered the dynamic population (~58%) transitioning between the three states. Binding of Mn^{2+} ions significantly shifted the transition towards higher access for I and H states, with a smaller population sampling the L state. The transition between L and I state resulted in a k_{dock} of 2.2 s^{-1} and k_{undock} of 2.6 s^{-1}, showing that the addition of Mn^{2+} resulted in faster k_{dock} and slower k_{undock} than in its absence (Figure 3.9, Top panel). The transition between the I and H states resulted in k_{dock} of ~2.4 s^{-1} and k_{undock} of 1.2 s^{-1}, slower than the rates obtained in the presence of Mg^{2+} only. The transition between the predominant L and H states resulted in a faster k_{dock} of ~2.2 s^{-1} and a much slower k_{undock} of ~3.5 s^{-1}, as compared to the condition with only Mg^{2+}. These results with Mn PEC 104 showed that while Mg^{2+} promotes dynamic transition between the three conformational states, the addition of Mn^{2+} is necessary to achieve the stable docked (H) conformation and bypasses the need for a transient intermediate (I) conformation to achieve docking.

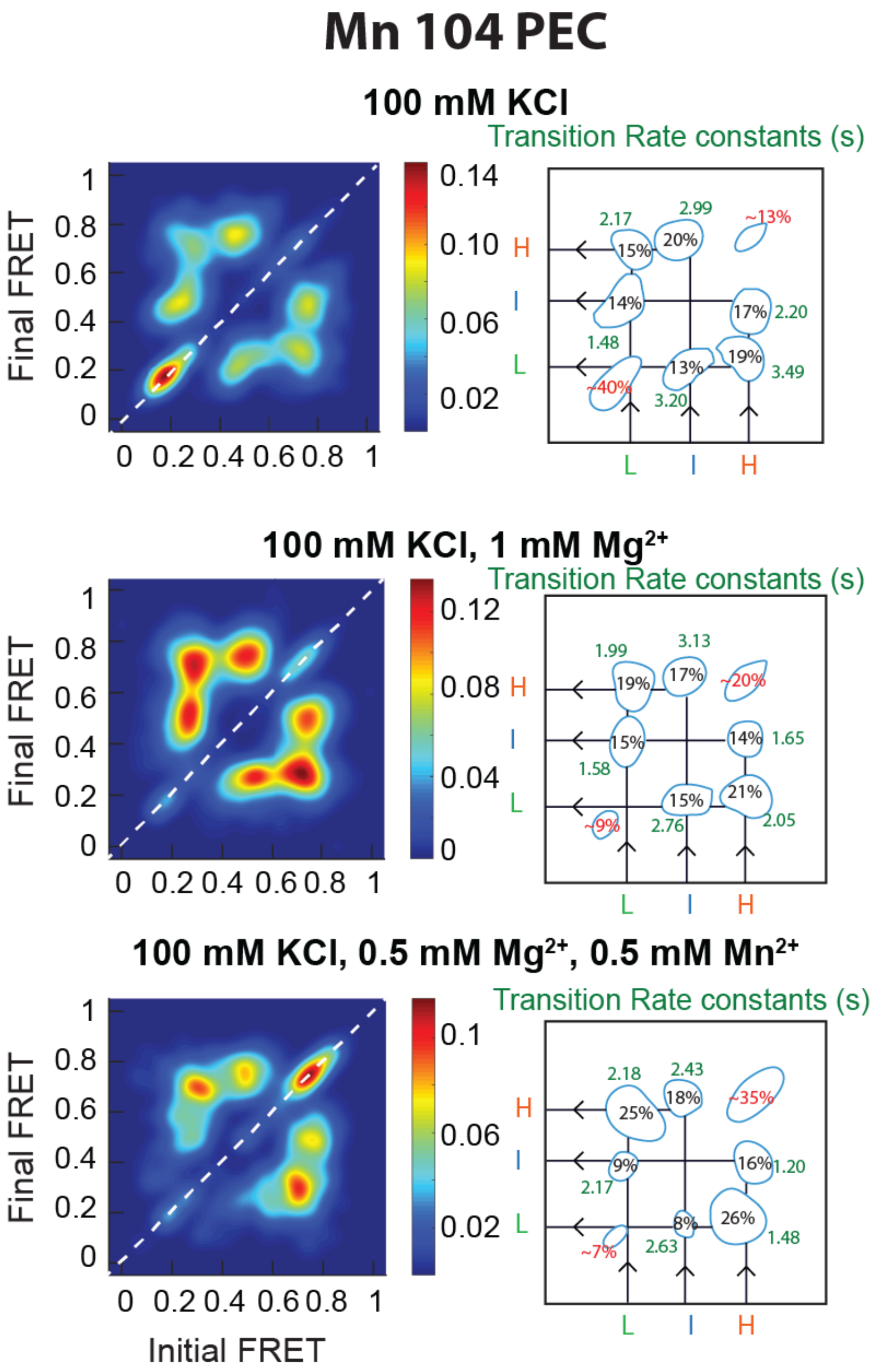


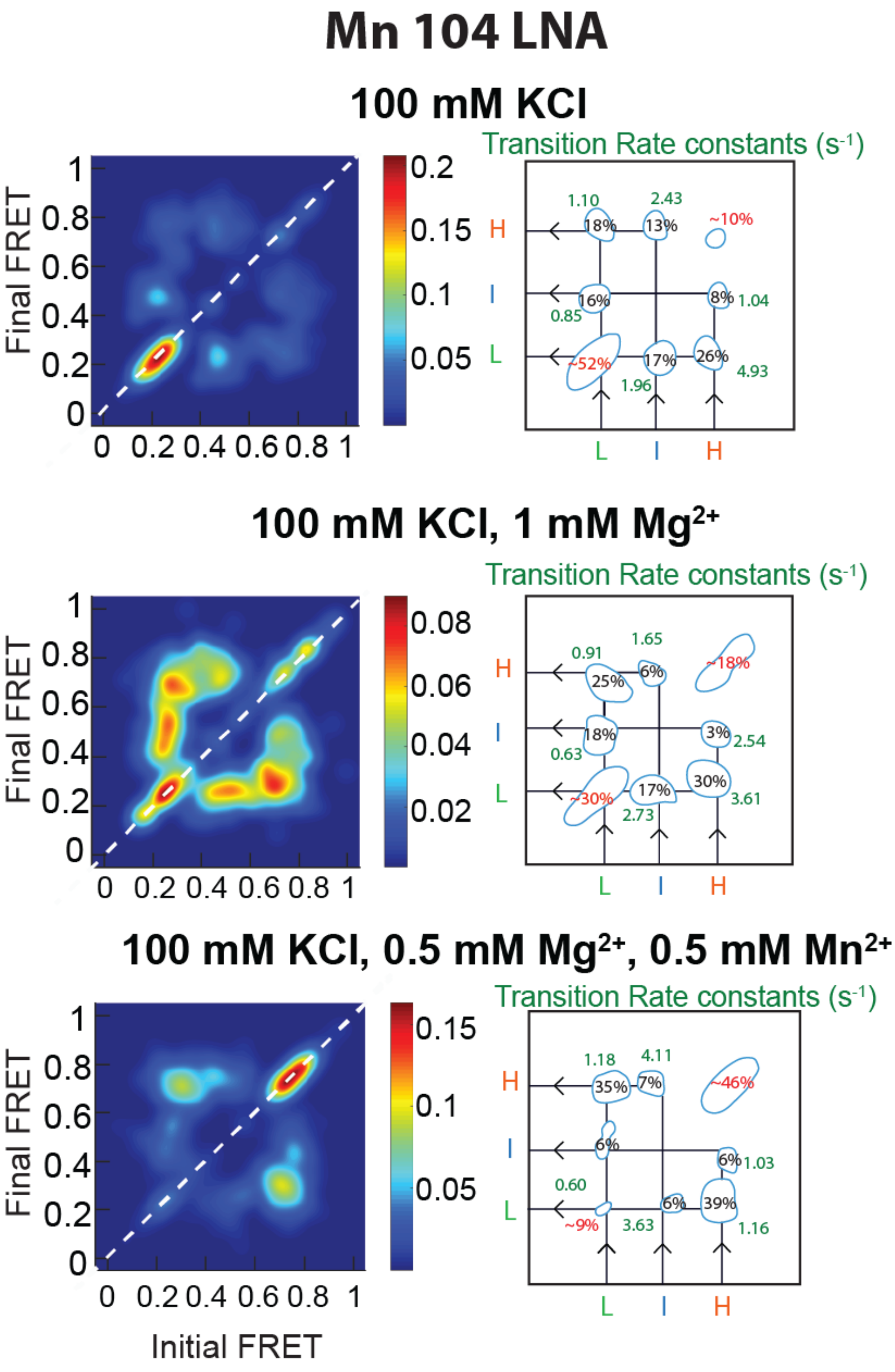
Figure 3.9 Transition between the FRET states in Mn 104 PEC show accessibility through an intermediate (I) state.

TODPs and their relative contributions and rate constants of all the transitions illustrated for each divalent ion conditions: (i) no divalent ions, (ii) 1 mM Mg^{2+} only, and (iii) 0.5 mM Mg^{2+} and 0.5 mM Mn^{2+}. The "on-diagonal" and "off-diagonal" contours represent static and dynamic

transitions, respectively. The color code on the TODPs indicates the fraction of each population, quantified on the right side.

To understand the role of RNAP in determining the conformational dynamics of the riboswitch, we estimated the static and dynamic populations and their transition kinetics for the Mn 104 LNA construct attached through a capture LNA instead of a PEC. At 100 mM KCl (Figure 3.10, Top panel), the riboswitch in Mn 104 LNA construct had its population predominantly in the stable undocked conformation (~52%) at ~0.2 FRET state and a smaller fraction (~38%) dynamically sampled between the three conformational states. The transition between L and I state resulted in a slow k_{dock} of 0.9 s^{-1} and a fast k_{undock} of 2 s^{-1}, showing that in the absence of RNAP, the riboswitch preferred to remain in the L state and undocked quickly while transitioning the intermediate conformation. The transition between I and H states, although least probable, resulted in k_{dock} of ~2.4 s^{-1} and k_{undock} of 1.0 s^{-1}. The slow undocking rate from H to I in the absence of RNAP showed that the transition via I state was more accessible when the riboswitch was present within a PEC. The transition between the L and H resulted in a slower k_{dock} of ~1.1 s^{-1} and a much faster k_{undock} of ~5.0 s^{-1}, agreeing with the fact that the docked H state is least accessible without any divalent ion. In the presence of 1 mM Mg^{2+} ions, the riboswitch attached to LNA sampled a lesser dynamic population (~50%) and a more stable undocked population (~30%) as compared to the riboswitch attached to the PEC (Figure 3.9 & 3.10, Middle panel). For the transitions between the three states, the transition through I states was lower in the absence of a PEC. The transition between L and I states resulted in a slower k_{dock} of 0.6 s^{-1} and a faster k_{undock} of 2.7 s^{-1}. These data show that, while the Mn PEC 104 favored docked transition though I, Mn LNA 104, in contrast favored less transition through the I state. The transition between I and H states remained the least probable, resulting in k_{dock} of ~1.7 s^{-1} and k_{undock} of 2.5 s^{-1}, showing that the addition of Mg^{2+} did not change the transition rate from undocked to docked conformation though I state. The transitions between the L and H were observed as the most probable, resulting in a slower k_{dock} of ~0.9 s^{-1} and a faster k_{undock} of ~3.6 s^{-1}. The absence of PEC in Mn 104 LNA resulted in a slower k_{dock} between L to H and a faster k_{undock} than the riboswitch within PEC, further supporting that the RNAP promoted the docking of the riboswitch. In the presence of Mn^{2+} ions (Figure 3.10, Bottom panel), the majority population predominantly remained in a stable docked population (~46%). Within the dynamic population, the addition of Mn^{2+} reduced the transition through I state. The transition between L and I state without PEC resulted in a much slower k_{dock} of 0.6 s^{-1} and a faster

k_{undock} of 3.6 s^{-1}, as compared to riboswitch with PEC. The transition between I and H states without PEC resulted in a faster k_{dock} of ~4.1 s^{-1} and slower k_{undock} of 1 s^{-1}, as compared to the riboswitch with PEC, showing that the addition of Mn^{2+} destabilized the transition through the I state in Mn 104 LNA, possibly due to higher accessibility to the ligand-binding core in the absence of RNAP. The transition between the L and H resulted in the most probable transition with $Mn^{2+,}$ resulting in a slow k_{dock} and k_{undock} of ~1.2 s^{-1} for both compared to riboswitch within the PEC. Overall, the Mn 104 LNA construct illustrated that, while the riboswitch predominantly transitions between two states (L and H), the accessibility through an intermediate conformational state (I) state was achieved more in the presence of PEC. The addition of Mn^{2+} to the riboswitch further promoted stable docked conformation in both constructs with or without RNAP. However, in the absence of PEC, the riboswitch in Mn 104 LNA adapted a close to two-state type transition, with less accessibility through I state.



Mn 104 LNA

100 mM KCl

100 mM KCl, 1 mM Mg^{2+}

100 mM KCl, 0.5 mM Mg^{2+}, 0.5 mM Mn^{2+}

Figure 3.10 Transition between the FRET states in Mn 104 LNA show lesser accessibility through an intermediate (I) state compared to Mn 104 PEC.

TODPs and their relative contributions and rate constants of all the transitions illustrated for each divalent ion conditions: (i) no divalent ions, (ii) 1 mM Mg^{2+} only, and (iii) 0.5 mM Mg^{2+} and 0.5 mM Mn^{2+}. The "on-diagonal" and "off-diagonal" contours represent static and dynamic

transitions, respectively. The color code on the TODPs indicates the fraction of each population, quantified on the right side.

To understand the role of formation and stabilization of P1.1, which was partially available within a PEC, but completely available when is RNAP is moved ten bases beyond the pause (G104), we looked at the static and dynamic transition from the TODP of Mn 114 LNA at different conditions of divalent ions (Figure 3.11). With a fully formed P1.1, the construct resembled an isolated riboswitch away from the influence of RNAP. In the absence of divalent ions (Figure 3.11, Top panel), The riboswitch predominantly adopted a stable undocked (L) conformation (~60%). Within the dynamic transitions, most populations were observed between L and I state, with a very slow k_{dock} of 0.3 s^{-1} and a very fast k_{undock} of 4.0 s^{-1}. The transition between I and H states resulted in a faster k_{dock} of ~2.9 s^{-1} and slower k_{undock} of ~2.0 s^{-1}, showing that I state is less accessible. As the riboswitch is well known to prefer undocked conformation in the absence of divalent ions[82], the transition between L and H states resulted in a slow k_{dock} of ~0.2 s^{-1} and a fast k_{undock} of ~8.8 s^{-1}. In the presence of 1 mM Mg^{2+} (Figure 3.11, Middle panel), the riboswitch sampled a similar number of dynamic population (~50%) as compared to the Mn 104 LNA (~50%). However, within the dynamic populations, the accessibility of I state was observed to be much lower than in Mn 104 LNA, showing that the intermediate conformational state (I) may be an outcome of partially formed P1.1 observed at pause 104. Even with three-state HMM fit, the riboswitch showed a two-state type transition predominantly between L and H state promoted by Mg^{2+} binding to the riboswitch. The transition between L and I stated showed a faster k_{dock} of 0.6 s^{-1} and a faster k_{undock} of 4.2 s^{-1} as compared to no divalent ions. In comparison with Mn 104 LNA, while the k_{dock} remained similar, k_{undock} increased, further showing the lower stability of I state when P1.1 is fully available. The transition between I and H states resulted in a slower k_{dock} of ~2.0 s^{-1} and a faster k_{undock} of ~2.5 s^{-1} as compared to no divalent ions, further supporting the less stability of the I state.

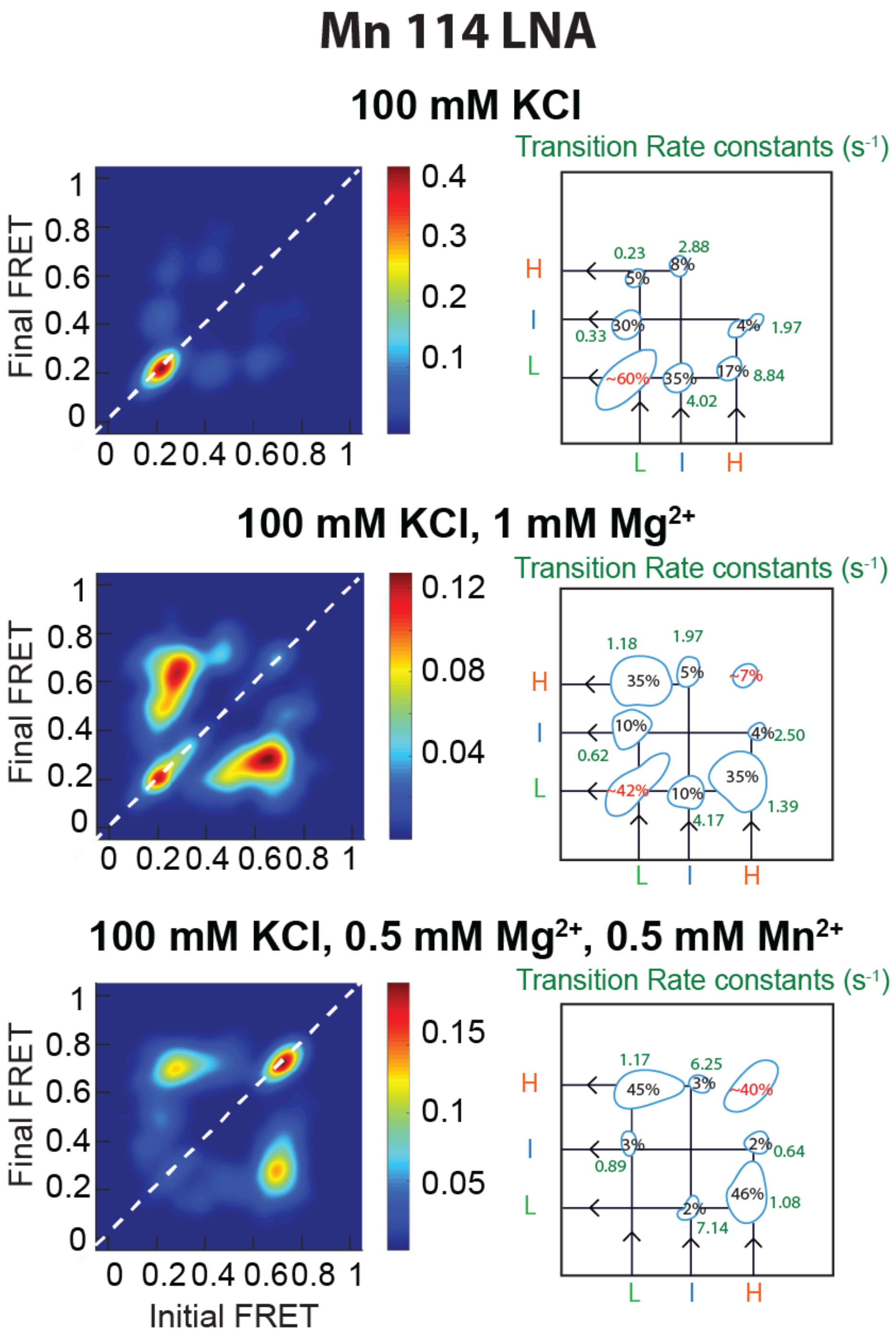


Figure 3.11 Transition between the FRET states in Mn 104 LNA show lesser accessibility through an intermediate (I) state compared to Mn 104 PEC.

TODPs and their relative contributions and rate constants of all the transitions illustrated for each divalent ion conditions: (i) no divalent ions, (ii) 1 mM Mg^{2+} only, and (iii) 0.5 mM Mg^{2+} and 0.5 mM Mn^{2+}. The “on-diagonal” and “off-diagonal” contours represent static and dynamic

transitions, respectively. The color code on the TODPs indicates the fraction of each population, quantified on the right side.

The most probable transition between L and H states resulted in a very faster k_{dock} of ~1.2 s^{-1} and slower k_{undock} of ~1.4 s^{-1}, showing that the addition of Mg^{2+} ions predominantly favored and stabilized the dynamic transitions between the L and H states, resembling with two-state type transition. In the presence of both Mg^{2+} and Mn^{2+} ions (Figure 3.11, Bottom panel), the majority of the population predominantly remained in a stable docked population (~40%), with the majority of dynamic transition happening between the L and H states. With the addition of Mn^{2+}, the transition between L and H resulted in a slow k_{dock} of 1.2 s^{-1} and k_{undock} of 1.1 s^{-1}, very similar to the condition in Mn 104 LNA.

Taking the kinetics data together showed that the riboswitch, when present within the PEC, adopts an intermediate conformational state (I) during the dynamic transitions between undocked (L) and docked (H) conformation. The probability for accessing the I state decreased in the absence of RNAP, showing that the I state conformation was more stabilized when the riboswitch was present within the exit channel of the RNAP. In Mn 104 PEC, the addition of Mn^{2+} to the riboswitch further destabilizes the transition through I state and promotes stable docking, showing that the riboswitch can attain docked conformation even with a partially formed P1.1 with two CC:GG base pairs available at the top of P1.1. In contrast, the Mn 104 LNA complex showed that the riboswitch without an RNAP has lesser access through I state and resulted in a predominant two-state (L↔H) type transition to access the docked state in the presence of Mn^{2+} when RNAP was absent. The accessibility of the I state further decreased for Mn 114 LNA, where the riboswitch can be expected to be away from the influence of RNAP. The least accessibility to I state in Mn 114 LNA showed that the I state conformation was transient and accessed only when riboswitch with partially available P1.1 was present within a PEC. However, when the riboswitch has fully available P1.1 (Mn 114 LNA), the accessibility through I state decreased.

3.3.5 Transcription factor NusA stabilizes the intermediate conformation for the riboswitch within a PEC

Transcription Factor NusA is known to stabilize hairpin regulated pauses and promote transcription termination by interacting with RNAP and nascent RNA transcript[139,140]. From our *in vitro* transcription assays, we observed that NusA enhanced pausing half-life at G104. To understand the mechanism for such enhancement in transcriptional pausing promoted by NusA, and its influence on the folding of riboswitch, we looked at the conformational change of the riboswitch in PEC (Mn 104 PEC) using smFRET in the presence of NusA (Figure 3.12A). In contrast to the smFRET data observed in the absence of the transcription factor, we observed more dynamic transition through the intermediate (I) state, whether in the absence or presence of divalent ions (Figure 3.12B).

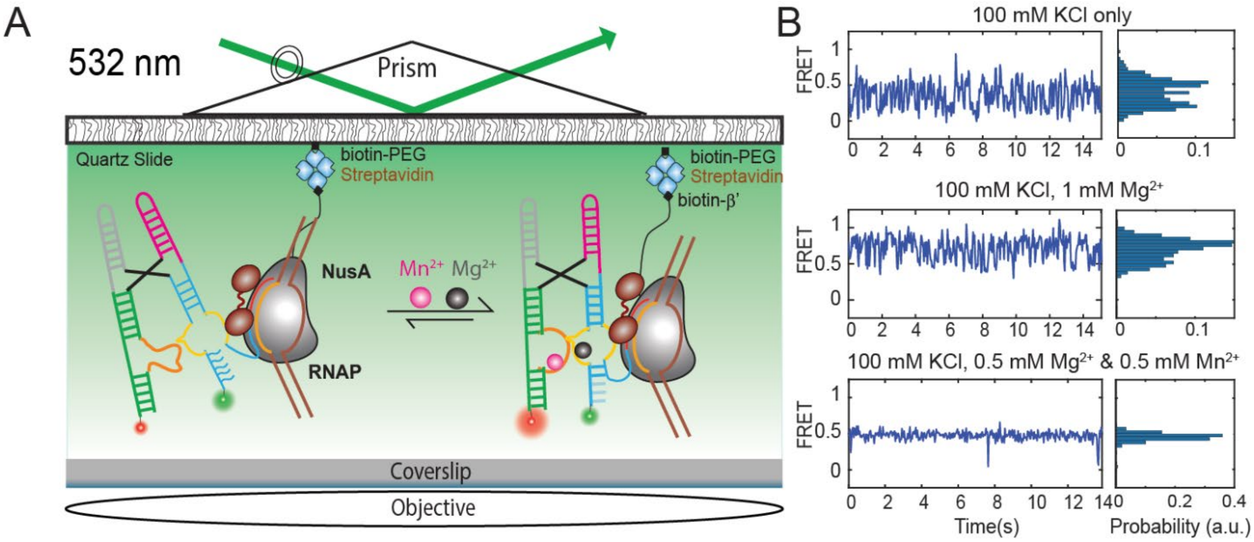


Figure 3.12 smFRET assay revealed the conformational dynamics of the Mn^{2+} sensing riboswitch immobilized through a PEC in the presence of NusA

A) Schematic of smFRET assay using pTIRFM. Dynamics of the riboswitch in a paused elongation complex (Mn-104 PEC) were studied in the presence of NusA, without and with the presence of Mn^{2+} added to the system. (B) Representative smFRET traces shown for the Mn-104 PEC with 350 nM NusA at different concentration of divalent ions (top to bottom) in the buffer: (i) no divalent ions, (ii) 1 mM Mg^{2+} only, and 0.5 mM Mg^{2+} and (iii) 0.5 mM Mn^{2+} respectively. Green, Cy3; Red, Cy5; Black, FRET. The occupancy of the signals over time in each represented trance are shown in the right side of the trace.

In the absence of any divalent ions (Figure 3.13A, Top panel), the cumulative histogram showed the majority of the population in the L state (~53%), a significant fraction of the population accessing the I state (~29%) and a small fraction accessing the H state (~19%). Looking at the transition kinetics between the conformational states observed from the TODP (Figure 3.13 B&C, Top panel) showed that ~35% of the total population sampled the stable docked conformation, and a small fraction (~5%) in stable intermediate conformation. The TODP further revealed that in contrast to earlier observations without NusA, the transition between L and H was least probable

(~16%), resulting in a k_{dock} of ~7.1 s^{-1} and k_{undock} of ~4.6 s^{-1}. The transition between L and I state resulted in a k_{dock} of 3.7 s^{-1} and k_{undock} of 4.4 s^{-1}. The most probable transition (~55%) between I and H states resulted in a k_{dock} of 7.1 s^{-1} and a k_{undock} of 3.2 s^{-1}. These rate constants showed that, in presence of NusA, the riboswitch undergoes fast docking and undocking predominantly between L and H states. In the presence of 1 mM Mg^{2+} ions (3.13A, Middle panel), the cumulative histogram showed an almost equal sampling of the population within the three states, with a significant increase in the I state population (~33%). The kinetics of transition between the conformational states observed from the TODP (3.13 B&C, Middle panel) further revealed the population sampled mostly dynamic transitions. However, a small fraction of stable undocked (~10%), stable intermediate (12%), and stable docked (~6%) present on the diagonal were sampled. The addition of Mg^{2+} showing that Mg^{2+} binding promotes both static and dynamic population within intermediate state. The transition between L and H resulted in a fast k_{dock} of ~6.7 s^{-1} and k_{undock} of ~3.9 s^{-1}, although least probable over other dynamic transitions. The transition between L and I state decreased with a slower k_{dock} of 4.8 s^{-1} and k_{undock} of 3.0 s^{-1}. The transition (~55%) between I and H states was observed to be even more probable (64%) and resulted in a slower k_{dock} of 5.3 s^{-1} and k_{undock} of 2.8 s^{-1}, as compared to the same condition in the absence of any divalent ions. With the addition of 0.5 mM Mn^{2+} with 0.5 mM Mg^{2+} ions (3.13A, Bottom panel), the cumulative histogram showed an increase in the I and H states population than without Mn^{2+}. The kinetics of transition between the conformational states observed from the TODP (3.13 B&C, Bottom panel) further revealed that ~25% of the static population were present in the stable intermediate state and ~30% in the stable docked state with the addition of Mn^{2+}. The majority of dynamic transitions (~64%) were observed between the I and H states, resulting in a slower k_{dock} of 4.2 s^{-1} and k_{undock} of 0.7 s^{-1}, showing that Mn^{2+} in the presence of NusA increased the accessibility to I state over L and H states and further stabilized the I state along with the H state.

Taken together, these data demonstrated the unique role of NusA in stabilizing an intermediate conformation (I) over the predominant undocked (L) and docked conformation (I). In absence of Mn^{2+}, NusA promotes the riboswitch to undergo a more dynamic transition through the I state. With the addition of Mn^{2+}, the riboswitch existed in both stable docked as well as static intermediate conformation. While the riboswitch attaining stable docked conformation was not surprising, the stabilization of the intermediate state with Mn^{2+} and NusA can be expected due to the countereffect of their action. NusA is well known to promote class I pausing and transcription termination[140]. In the present case, NusA promotes stability of the intermediate conformation over the stable docked conformation, suggesting that NusA can influence riboswitch structure during pausing to promote transcription termination.

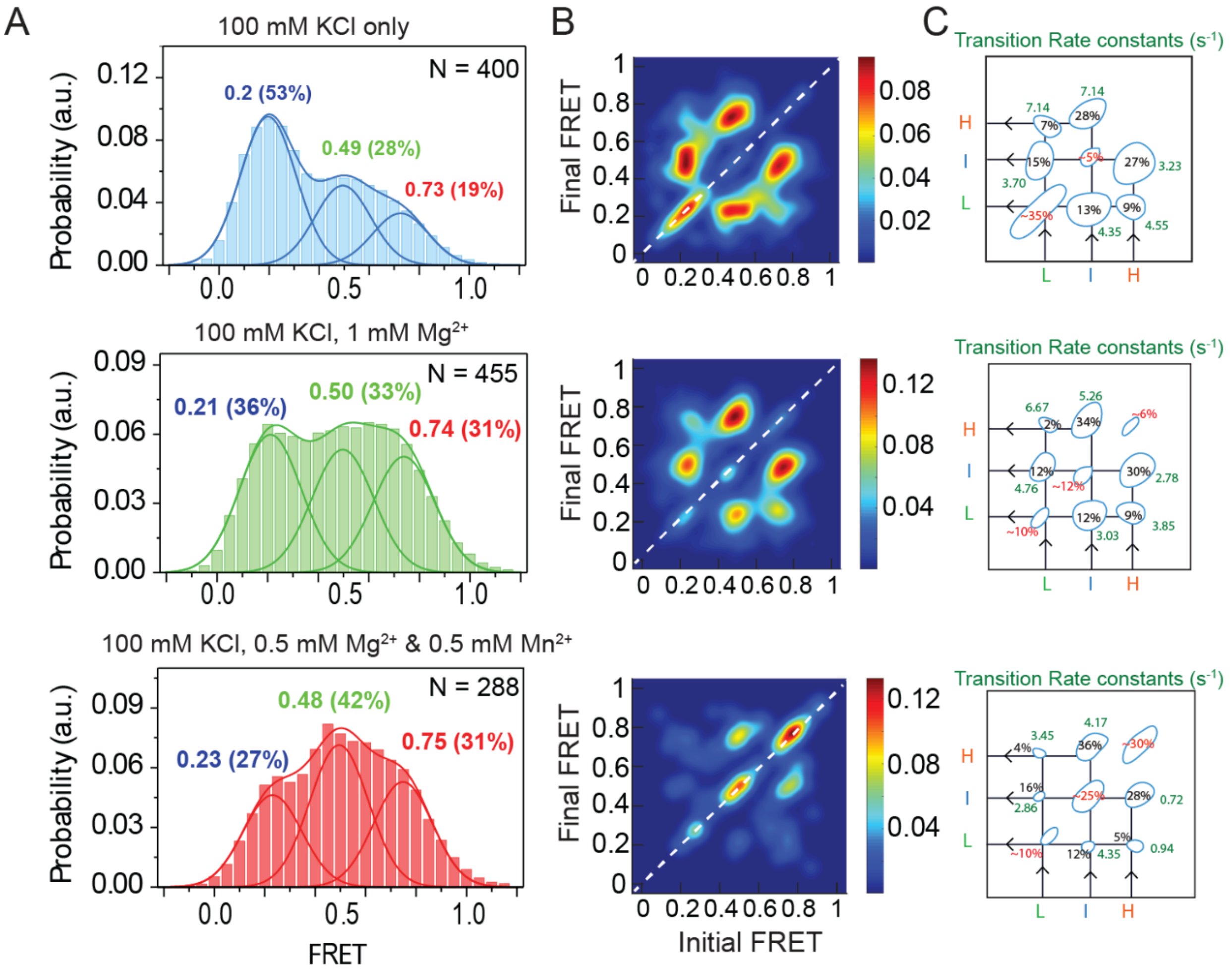


Figure 3.13 Mn^{2+} ion in the presence of NusA promotes stability of intermediate conformation over the undocked and docked conformation

(A) Population of FRET histograms for each divalent ion conditions: (i) no divalent ions, (ii) 1 mM Mg^{2+} only, and (iii) 0.5 mM Mn^{2+} and 0.5 mM Mg^{2+}, recorded for Mn-104 PEC in the presence of NusA. Each histogram is fitted with 3 Gaussian peaks to represent their predominant FRET population densities. (B-C) TODPs and their relative contributions and rate constants of all the transitions illustrated for each divalent ion conditions ((i) no divalent ions, (ii) 1 mM Mg^{2+} only, and 0.5 mM Mg^{2+} and (iii) 0.5 mM Mn^{2+}), recorded for Mn-104 PEC in the presence of NusA. The “on-diagonal” and “off-diagonal” contours represent static and dynamic transitions, respectively. The color code on the TODPs indicates a fraction of each population, quantified on the right side.

3.4 Discussion

Earlier studies applying biophysical and biochemical assays to co-transcriptional folding riboswitches have shown that pausing of RNAP during transcription allows time for ligand binding and riboswitch folding[23,135,146,151]. However, the mechanism of riboswitch folding at pause is rarely addressed. Our present work provides insights into the role of a class I pause in determining the conformational dynamics of a transcriptional riboswitch. Using a combination of biochemical assays and single-molecule assays, we demonstrated the role of transcriptional pausing on the co-transcriptional folding and conformational dynamics of a Mn^{2+} sensing riboswitch from *L. lactis*. In light of the previous works on folding of riboswitches within a paused elongation complex[23,57], our work provides mechanistic insights into the dynamics of a Mn^{2+} sensing riboswitch and the role of a ligand-dependent pause on determining the transcriptional outcome.

Previous studies have reported the structure with conformational dynamics and the ligand sensing mechanism for the Mn^{2+} ion sensing riboswitch aptamer[82,124]. However, riboswitch folding *in vivo* occurs co-transcriptionally to determine the fate of transcription, where RNAP and transcription events like pausing influence the folding of the riboswitch. Using the Mn^{2+} sensing riboswitch within a paused elongation complex, we studied the conformational dynamics as well as formation and stabilization of the key switch helix P1.1, which could determine the outcome of transcription. Using biochemical assay, we identified a transcription pause site at the terminator region of the Mn^{2+} sensing riboswitch. We found that the half-life of the pause increased with the presence of Mn^{2+} ions and NusA during the transcription. A nearly four-times increase in the pause half-life in presence of Mn^{2+} showed that the addition of ligand during transcription allowed the RNAP to pause longer to help the riboswitch in stabilizing the key switch helix P1.1 between riboswitch aptamer and the terminator hairpin (Figure 3.2). Upon mutation of the P1.1 to destabilize its formation, we observed no influence of Mn^{2+} addition on pausing. In addition to these observations, an increase in the pause half-life with the addition of NusA further established the nature of pause as a Class I pause, which is known to be stabilized by the formation of a "pausing hairpin" in the transcribed RNA[136,152].

Using a combination of different single-molecule assays, we showed that pausing of RNAP alongside ligand binding has a vital role in determining the fate of co-transcriptional folding and conformational dynamics of a riboswitch. Our SiM-KARTS assay with the co-transcriptionally folded riboswitch in a PEC where RNAP is halted at position 104 demonstrated that the presence of RNAP promoted folding of the riboswitch structure through the formation of the partially available P1.1. Such folding can be expected due to the steric interaction of the large RNAP complex on the transcribing riboswitch at the RNAP exit channel (Figure 3.3). We further showed that the addition of Mn^{2+} further promotes the stabilization of the partially formed P1.1 to attain riboswitch folding (Figure 3.3). However, the effect of Mn^{2+} binding to the riboswitch leading to its folding and stabilization is overshadowed by the influence of the RNAP. Previous studies with Mn^{2+} sensing riboswitches have shown that the formation of P1.1 with the binding of Mn^{2+} to the ligand-binding core L1 is essential for preventing the formation of the terminator hairpin and consequent upregulation of the gene[82,106]. Using SiM-KARTS assay, we showed that binding of Mn^{2+} to the riboswitch could promote docking the riboswitch even with a partially transcribed P1.1 helix.

Using smFRET, we further revealed the conformational switching of the riboswitch within a PEC and its effect with the addition of Mn^{2+} in contrast with the earlier observations with aptamer. We identified a previously unknown intermediate conformation (I) between the previously known docked and undocked conformations (Figure 3.6). Under the sterically constrained environment of the RNAP exit channel, the riboswitch was observed to adopt a partially folded conformational intermediate. Our kinetic analysis from TODPs showed faster transition through the intermediate state than the predominant low and high FRET state conformations in the absence of Mn^{2+} (Figure 3.9). However, transition through the intermediate state was reduced as the riboswitch within a PEC binds to Mn^{2+} ions (Figure 3.9). In contrast with the riboswitch within the PEC (Mn PEC 104), the intermediate state becomes less accessible as the RNAP moves beyond pause (Mn LNA 114) and riboswitch moves away from the steric constrain of the RNAP exit channel (Figure 3.11). We further observed that the occurrence of the intermediate FRET conformation increased with NusA and further stabilized with the addition of Mn^{2+} (Figure 3.12). Since NusA is known to stabilize class I pausing and promote transcription termination[140], which for the manganese riboswitch can only be achieved by disruption of P1.1 to form terminator hairpin, our data herein support that NusA destabilizes P1.1 to oppose the effects of Mn^{2+} by instead stabilizing an intermediate conformation over the stable docked conformation. These results support previous observations that pausing during transcription elongation acts as a decision point to determine the folding of transcribing RNA[137].

From our previous understanding of the Mn^{2+} riboswitch and developments from the present study, we propose a model for the conformational switching of the riboswitch present within a PEC (Figure 3.14). In the absence of any divalent ion in the solution, the riboswitch emerging from the RNAP adapts an open/undocked conformation. However, within the steric environment of the PEC, the presence of RNAP promotes folding of the riboswitch by allowing the formation of P1.1. Under these conditions, the riboswitch can adopt a short-lived intermediate conformation. With the addition of Mg^{2+}, the riboswitch conformations become more dynamic in nature with more access through the intermediate conformation to attain the docked and undocked states. During the folding of the riboswitch when Mg^{2+} is bound to one of the ligand-binding cores, the riboswitch starts forming docked conformation by stabilizing the partially available P1.1 to allow the riboswitch to sense Mn^{2+} ion. When Mn^{2+} is bound to the riboswitch, it stabilizes the ligand-binding core, and thus promote stabilization of even the partially formed P1.1 with two base pairs to attain a docked conformation. Our studies show that the pausing of RNAP can be essential in allowing time for the riboswitch to bind divalent ions necessary for stabilization of P1.1 to upregulate transcription. However, during a pause, if the riboswitch does not bind Mn^{2+}, it keeps sampling the conformation states that are unable to stabilize P1.1. In such conditions, as transcription elongation proceeds beyond pause, the P1.1 region helps in forming terminator hairpin and downregulation of associated genes.

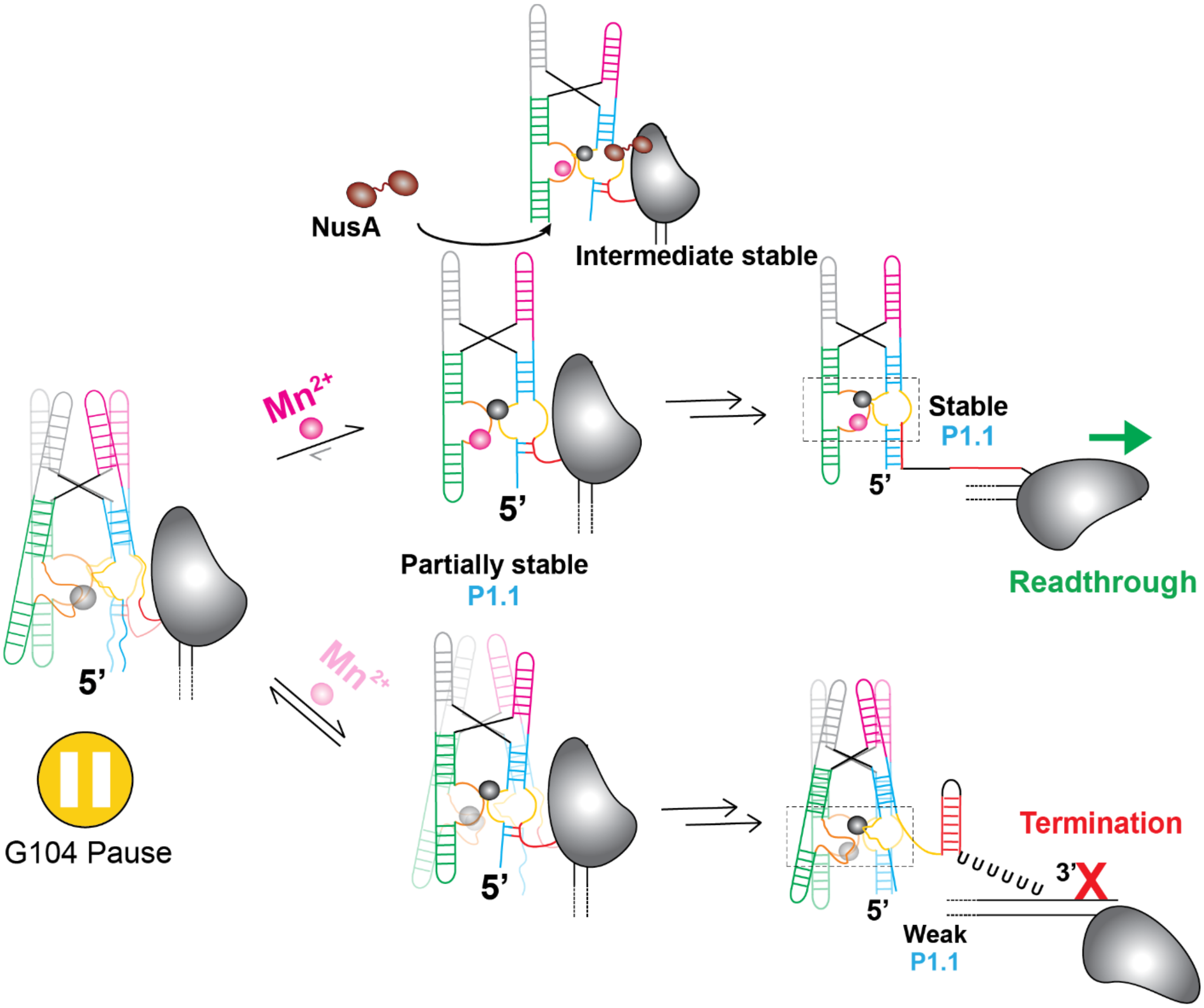


Figure 3.14 Proposed model for illustrating the role of G104 pause, paused RNAP and Mn^{2+} binding in determining the conformational state of the riboswitch for ultimately deciding the fate of gene expression.

The 4WJ of *L. lactis* riboswitch paused at G105 samples predominantly dynamic population due to partially formed P1.1 stem region as well as loosely stabilized ligand-binding core (L1 & L3) in the absence of Mn^{2+} ion. In the presence of a physiological concentration of Mg^{2+} (1 mM), the riboswitch pre organized with ligand binding site poised for binding Mn^{2+} ion. Pausing of RNAP allows time for the riboswitch to decide on binding Mn^{2+} necessary for upregulation. When Mn^{2+} ion is present at sub-millimolar concentration, it binds to L3 and stabilizes the ligand-binding core, as starts stabilizing partially available P1.1 to promote the docking of riboswitch. This prevents the formation of a terminator and signals the RNAP to move forward for upregulating transcription. Addition of NusA promotes stable intermediate conformation that act opposite to the effect of Mn^{2+}. On the other hand, during pausing, if the riboswitch does not bind to Mn^{2+} ion, it is unable to stabilize P1.1 and allow the formation of termination hairpin that ultimately results in down-regulation of transcription.

Riboswitches regulating transcription are known to be kinetically controlled, where the speed of transcription determine the ligand binding efficiency and structural change[57,144,153]. Mechanism like pausing of elongating RNAP at strategic positions plays an important role to balance the speed of riboswitch folding and transcription elongation[146]. Studies with coenzyme B12 binding aptamer in the *btuB* gene have shown that transcriptional pausing is a strategic mechanism that coordinates riboswitch folding and conformational rearrangement to control the outcome of the expression platform[146]. PreQ$_1$ sensing riboswitch studied in a PEC have shown that conformational dynamics of riboswitch present in the exit channel of the RNAP is significantly influenced by the steric interaction due to RNAP[23]. Our work adding on to these studies shows that pausing helps in determining key conformational states of a riboswitch that may ultimately determine the fate of the transcription regulation. While the addition of ligand (Mn^{2+}) is a key determinant for conformational states, the occurrence of a previously unknown conformation state (I) showed that pausing of RNAP close to the aptamer strongly influenced the riboswitch structure and co-transcriptional folding of the riboswitch. Recent smFRET studies with Nickel-Cobalt (NiCo) sensing riboswitches have shown that the occurrence of such an intermediate state in a 4WJ may be due to the pre-folded intermediate that reorganizes the ligand-binding pockets to allow selective binding of divalent ions[154]. In contrast, with our observations on the Mn^{2+} sensing riboswitch, the occurrence of the intermediate state and transitions through it can be due to a combination of effects from pre-organization of the riboswitch ligand-binding core as well as the partially available P1.1 under the steric environment of the RNAP exit channel. Our work outlines the mechanism behind such pre-organization of the riboswitch and its influence under the addition of Mn^{2+} within a PEC, providing new insights into the previous understanding of transcriptional pausing and riboswitch folding for fine-tuning of transcription regulation in bacteria. While more structural insights are needed to characterize the riboswitch conformations within a PEC to achieve deeper insights into the intermediate conformation, the present work demonstrates how the Mn^{2+} ion alongside RNAP and transcription factor NusA control riboswitch folding in both agonistic and antagonistic ways. Our results overall provide new mechanistic insights into our understanding of riboswitch regulated gene expression, which can open the door for leveraging these systems for anti-microbial drug targets[18,155-157].

3.5 Methods and Materials

3.5.1 *In vitro* transcription assays

Halted complexes (PEC 104) were prepared in transcription buffer (20 mM Tris-HCl, pH 8.0, 20 mM NaCl, 20 mM $MgCl_2$, 14 mM 2-mercaptoethanol, 0.1 mM EDTA) containing 25 µM ATP/CTP mix, 50 nM α^{32}P-GTP (3000 Ci/mmol), 10 µM ApU dinucleotide primer (Trilink), and 50 nM DNA template. *E. coli* RNAP holoenzyme (New England Biolabs) was added to 100 nM, and the mixture was incubated for 10 min at 37 °C. The sample was passed through G50 column to remove any free nucleotides. To complete the transcription reaction all four rNTPs were added concomitantly with heparin (450 µg/mL) to prevent the re-initiation of transcription. Time pausing experiments were performed using 10 µM rNTPs without and with 0.5 mM $MnCl_2$. The mixture was incubated at 37 °C, and reaction aliquots were quenched at the desired times into an equal volume of loading buffer (95% formamide, 1 mM EDTA, 0.1% SDS, 0.2% bromophenol blue, 0.2% xylene cyanol). Sequencing ladders were prepared by combining the halted complex with a chase solution containing 250 µM of each rNTP, in addition to one 3'-OMe rNTP (at 25 µM for 3'-OMe GTP and 15 µM for 3'-OMe ATP, UTP and CTP). Reaction aliquots were denatured before loading 5 µL of each onto a denaturing 8 M urea, 6% polyacrylamide sequencing gel. The gel was dried and exposed to a phosphor screen (typically overnight), which was then scanned on a Typhoon Phosphor Imager (GE Healthcare).

3.5.2 Transcription data analysis

To determine the T_{50} for regulation of termination by fluoride, the relative intensity of the full-length product was divided by the total amount of RNA transcript (full-length + terminated product) for each fluoride concentration. The resulted percentages of readthrough were plotted against the fluoride concentration using equation: $Y=Y_0 + ((M1\times X)/(T_{50}+X))$, where Y_0 is the value of Y at X_{min} (no ligand condition) and $M1 = Y_{max} - Y_{min}$. The half-life of transcriptional pausing was determined by calculating the fraction of each RNA pause species compared with the total amount of RNA for each time point, which was analyzed with pseudo-first-order kinetics to extract the half-life[158]. For each determination we have subtracted the background signal. Error bars in transcription quantification represent the standard deviation of the mean from at least two independent replicates.

3.5.3 Preparation of fluorescently labeled nascent transcripts for SiM-KARTS assay

In vitro transcription reactions were performed in two steps to allow the specific incorporation of Cy3 at the 5'end of the RNA sequence. Transcription reactions were performed in the same transcription buffer described above. Transcription reactions were initiated by adding 100 µM AU-Cy3 dinucleotide and 25 µM ATP/CTP/GTP nucleotides at 37°C for 10 min, thus yielding a fluorescent halted complex through biotin-streptavidin roadblock[137]. The resulted halted complex was immobilized to the microscope slide surface using the DNA templates contain a biotin molecule at the 5'-end of the antisense DNA strand. $MnCl_2$ (when present) was always added during the transcription of the riboswitch at a final concentration of 0.5 mM.

3.5.4 Preparation of fluorescently labeled PEC for smFRET

Two fluorescently labeled strands for the Mn^{2+} sensing riboswitch from *L. lactis* were purchased from GE Darmacon. The riboswitch was annealed with the template and non-template DNA for a total concentration of 1 µM. During annealing, 1 µM of two RNA strands with two DNA templates were mixed (1:1:1:1 ratio) using a nucleic acid scaffold buffer (50 mM Tris-HCl, pH 7.5 and 100 mM KCl). The solution was heated at 90°C for 2 minutes followed by incubation at 37°C for 10 minutes and then at room temperature for 10 minutes before storing at 4°C. This annealed complex was stored at -20°C for future experiments.

For reconstitution of PEC, the annealed complex was assembled with biotinylated RNAP (1:1 ratio) for a total concentration of 50 nM using an assembly buffer (50 mM Tris-HCl, pH 7.5, 100 mM KCl and 1mM $MgCl_2$). The mixture was incubated at 37°C for 15 minutes and stored at 4°C for the duration of single-molecule experiments. 50 pM of the reconstituted PEC was added to the slide for single-molecule imaging.

For LNA based immobilization of riboswitch to the slide surface, the two strands of fluorescently labeled RNA were annealed with an LNA capture strand containing biotin (1:1:1) for a total concentration of 1uM. The mixture) using a nucleic acid scaffold buffer (50 mM Tris-HCl, pH 7.5 and 100 mM KCl). The solution was heated at 90°C for 2 minutes followed by incubation at 37°C for 10 minutes and then at room temperature for 10 minutes before storing at 4°C. This annealed complex of 50 pM was used for smFRET experiments and the stock solution was stored at -20°C for future experiments.

3.5.5 Single-molecule experiments

Single-molecule experiments were performed using a prism-based TIRF (pTIRF) microscope. All movies SiM-KARTS were collected at 100 ms time resolution. All movies for SMFRET were collected at 60 ms time resolution. PEG-passivated quartz slides with a microfluidic channel containing inlet and outlet ports for buffer exchange were assembled as described in the previous works[44]. For real-time transcription assay, the inlet consists of a cut pipet tip acting as a reservoir, and the outlet is connected to a 3 mL syringe to pull-in the solution from the reservoir to the microscope slide. The surface of the microfluidic channel was coated with streptavidin (0.2 mg/mL) for 10-15 min prior to flowing RNA transcript. In the case of PEC, the fluorescent PECs were directly injected into the channel surface using the biotin-streptavidin roadblock for immobilization. For SiM-KARTS experiments, the RNA molecules were diluted in SiM-KARTS buffer (80 mM HEPES-KOH pH 7.5, 300 mM KCl, 5 mM $MgCl_2$). For smFRET experiments, the nascent RNA molecules were diluted in smFRET imaging buffer (20 mM HEPES-KOH pH 7.5, 50 mM KCl). We included an enzymatic oxygen scavenging system (OSS) containing 50 nM protocatechuate-3,4-dioxygenase and 5 mM protocatechuic acid in 1x buffer to prolong the longevity of the fluorophores. In addition, 2 mM Trolox (6-Hydroxy-2,5,7,8-tetramethylchromane-2-carboxylic acid) was included in the imaging buffer to suppress photoblinking of the dyes. For SiM-KARTS assay, the raw movies were collected for 10-15 mins per movie with direct green (532 nm) and red (638 nm) lasers excitation. For smFRET assay, the raw movies were collected for 2-3 mins (at 60 ms time resolution) per movie with direct green (532 nm).

3.5.6 Data acquisition and analysis for SiM-KARTS assay

Locations of molecules and fluorophore intensity traces for each molecule were extracted from raw movie files using custom-built MATLAB codes. Traces were manually selected for further analysis using the following criteria: single-step photobleaching of Cy3, ≥ 2 spikes of Cy5 fluorescence of more than 2-fold the background intensity. Traces showing binding events were idealized using a two-states (bound and unbound) model using a segmental *k*-means algorithm in QuB[44,114]. From the idealized traces, dwell times of NusA or the SiM-KARTS probe in the bound (τ_{bound}) and the unbound ($\tau_{unbound}$) states were obtained. Cumulative plots of bound and unbound dwell-time distributions were plotted and fitted in Origin lab with single-exponential or double-exponential functions to obtains the lifetimes in the bound and unbound states. The dissociation rate constants (k_{off}) were calculated as the inverse of the τ_{bound}, whereas the association rate constants (k_{on}) were calculated by dividing the inverse of the $\tau_{unbound}$ by the concentration of SiM-KARTS probe used during the data collection[65].

3.5.7 Data acquisition and analysis for smFRET assay

Locations of molecules and fluorophore intensity traces for each molecule were extracted from raw movie files using custom-built MATLAB codes. Single-molecule traces were then visualized using Matlab and only those with a minimum combined intensity (Cy3 + Cy5 intensity) of 300 a.u., showing single-step photobleaching of the dyes, a signal-to-noise ratio of >3, and longer than 6 s were selected for further analysis. Selected traces were then background-subtracted to correct for cross-talk and (minimal) bleed-through. We calculated the FRET ratio as $I_A/(I_A+I_D)$, where I_A and I_D are the background-corrected intensities of the acceptor (Cy5) and donor (Cy3), respectively. FRET histograms were made using the first 100 frames of all traces in a given condition and fit with a sum of Gaussians using OriginPro 8.5. For kinetic analysis, traces were idealized with a three-state model corresponding to undocked (low-FRET), intermediate docked (mid-FRET) and docked (high-FRET) states using the segmental k-means algorithm in QuB software as previously described[114,129]. Cumulative dwell-time histograms were plotted from all extracted dwell times and fit with single- or double-exponential functions using OriginPro 8.5 to obtain the lifetimes in the undocked (τ_{undock}) and docked (τ_{dock}) states. Rate constants of docking and undocking were then calculated as $k_{dock} = 1/\tau_{undock}$ and $k_{undock} = 1/\tau_{dock}$. For the double-exponential fits, kinetics were calculated similarly using both the short and long dwell lifetimes to obtain the fast and slow rate constants, respectively. The idealized smFRET traces were used for creating transition occupancy density plots (TODPs), which show the fraction of traces/molecules that exhibit a given type of transition at least once[114]. In TODPs, dynamic traces showing a FRET transition (regardless of the number of transitions in that trace) and static traces (with no transitions over the entire trace) are weighted equally, avoiding over-representation of the traces with fast transitions.

3.5.8 RNA and DNA constructs for smFRET assays

Construct	Sequence

Mn riboswitch RNA strand 1	5’/Cy3/UCAAAGGGGAGUAGCGUCGGUAAGACCGAA ACAAAGUCGUCAAUUCGUGAGAU/3’
Mn riboswitch RNA strand 2	5’/Cy5/UCUCACCGGCUUUGUUGACAUACUUAUGUA UGUUUAGCAAGACCUUUGCCAG/3’
10 base RNA extension at 3’ of strand 2 for Mn 114 LNA	5Phos/CUGGCAGAGG - 3'
Template DNA for transcription bubble	CCAGTC ATG CAG GAC TGG CAA ACC TTC TCTAAG TCT C
Non-template DNA for transcription bubble	GAGACT TAG AGA AGG AAA TGG TTA CCT GCATGA CTG G
LNA capture strand for Mn 104 LNA	/5Biosg/TTTTT+CT+GG+CA+AA
LNA capture strand for Mn 104 LNA	/5Biosg/TTTTTCC+TC+TGCC+AGC

Chapter 4. Precise Tuning of Bacterial Translation Initiation by Non-Equilibrium 5'-UTR Unfolding Observed in Single mRNAs[4]

4.1 Overview

Noncoding, structured 5'- untranslated regions (5'-UTRs) of messenger RNAs (mRNAs) control translation efficiency by forming structures that can either recruit or repel the ribosome. Here we exploit a bacterial, preQ$_1$-sensing translational riboswitch to probe how binding of a small ligand control binding of the bacterial ribosome to the Shine-Dalgarno (SD) sequence. Integrating single-molecule fluorescence microscopy with mutational analyses, we find that the stability of 30S ribosomal subunit binding is inversely correlated with the free energy needed to unfold the 5'-UTR during mRNA accommodation from the standby site to the binding cleft. Ligand binding stabilizes 5'-UTR structure to both antagonize 30S recruitment and accelerate 30S dissociation. The resulting model of rapid standby site exploration followed by gated non-equilibrium unfolding of the 5'-UTR during accommodation provides a mechanistic understanding of translation efficiency.

4.2 Introduction

5'-untranslated regions (5′-UTRs) of messenger RNAs (mRNA) are essential for regulating protein expression in all cells[159-163]. Direct interaction between the smaller or 30S ribosomal subunit and the 5'-UTR allows modulation of initiation as the rate-limiting step of bacterial mRNA translation[164-167]. The early stage of initiation involves two steps as major contributors to overall translation efficiency: reversible binding of the mRNA to a loosely defined "standby site", followed by accommodation into the mRNA binding cleft of the 30S subunit[168-170]. The accommodated 30S ribosomal binding site (RBS) of the mRNA stretches from nucleotide −18 to +10 relative to the start site (position +1) and often encompasses a purine-rich region around

---

[4] Some contents of this chapter are adapted from a manuscript of the same title under preparation, by Sujay Ray, Shiba S. Dandpat, Surajit Chatterjee and Nils G. Walter. S.R. and N.G.W. designed the study. S.R, S.D, S.C performed experiments and analyzed the data the paper. S.R, S.D, S.C, and N.G.W. wrote the manuscript.

positions -7 to -4 known as the Shine Dalgarno (SD) sequence[171]. This sequence directly engages with the 3'-terminus of the 16S ribosomal RNA (rRNA) of the 30S subunit – the so-called anti-SD sequence – through base pairing, requiring the RBS to be void of secondary structure and result in an active translation initiation complex. Translation initiation thus requires unfolding of mRNA structure near the start codon, imposing a structure-dependent energetic penalty on translation efficiency[166,172-176]. Consequently, transcriptome-wide studies have shown that an increased translation efficiency of a given mRNA generally correlates with its reduced propensity to form secondary structures near the RBS[174,175,177,178].

While the absence of local secondary structure permits efficient translation initiation on mRNAs that lack an SD sequence[179], structured RNA motifs such as riboswitches embedded in a 5'-UTR require an SD sequence for efficient gene expression, enabling dynamic regulation[96,99,155,180-183]. Primarily found in bacteria, riboswitches are cis-regulatory elements typically controlling either transcription or translation[99,155,183]. Following ligand binding to its aptamer domain, a typical translational riboswitch changes the secondary structure of the downstream expression platform to sequester the SD sequence, thus repressing translation initiation[184].

The extent of riboswitch-mediated translational control varies across many riboswitch classes discovered to date. In general, larger and more complex riboswitches may entirely abrogate SD and/or start codon access, leading to switch-like ON/OFF behavior[185,186]. In contrast, smaller riboswitches, as exemplified by the H-type pseudoknot of the class-I preQ$_1$ (or 7-aminomethyl-7-deazaguanine) riboswitch from *Thermoanaerobacter tengcongensis* (*Tte*), affect gene expression through only partial SD sequestration (Figure 4.1)[43,149,187-189]. How binding of a small ligand, which only yields little thermodynamic free energy (~10 kcal/mol)[43,190,191], regulates mRNA accessibility by the much larger bacterial 30S subunit has remained a conundrum. Transcriptome-wide SHAPE-MaP analysis of translation initiation supports a model wherein the mRNA first binds to a standby site on the 30S subunit[166], followed by transient unfolding of 5'-UTR structure during accommodation, imposing a non-equilibrium energetic penalty on translation initiation[168,169]. These findings suggest that the relative kinetics of riboswitch folding versus 30S binding may play a role in regulation[192,193], however, the mechanism of such a kinetically controlled process has not been observed directly.

In this study, we develop a Single-Molecule Kinetic Analysis of Ribosome Binding (SiM-KARB) assay, based on the previously developed SiM-KARTS assay[37,65], using a labeled bacterial 30S ribosomal subunit and perform mutational analyses to probe the mechanistic basis for the control of translation initiation by the *Tte* preQ$_1$ riboswitch. Repetitive binding and dissociation of the 30S subunit to the riboswitch-hosting mRNA (henceforth referred to as "R-mRNA") reveals short binding events that reflect non-specific "standby site" interactions, whereas significantly longer binding events represent "initiation-active" interactions where R-mRNA becomes partially unfolded and accommodated into the mRNA binding cleft of the 30S subunit. Moreover, while both standby and initiation-active binding are profoundly affected by preQ$_1$, SD:anti-SD complementarity and aptamer proximity each have more differentiated impacts. Taken together, our data support a model wherein the stability of 30S binding is inversely related to the non-equilibrium free energy of aptamer unfolding during mRNA accommodation, enabling the ligand to reduce both 30S binding speed and kinetic stability as two distinguishable mechanisms for modulating translation efficiency**.**

4.3 Results

4.3.1 Riboswitch ligand modulates 30S binding to the 5'-UTR of the mRNA

Prior crystallographic and single-molecule FRET studies of its isolated aptamer domain have suggested that the *Tte* preQ$_1$ riboswitch controls translation initiation by the ligand-induced sequestration of two nucleotides shared between the P2 helix of the pseudoknot and the SD sequence[22,43,189,194] (Figure 4.1A, 4.1B). The 30S subunit (Figure 4.1C), however, occupies a total of ~30 nucleotides in the 5'-UTR that unfold when stably accommodated into the mRNA binding cleft and are thus considered part of the expression platform (Figure 4.1D). We transcribed a *Tte* gene-1564 sequence *in vitro* that includes this entire expression platform under the influence of the preQ$_1$ riboswitch (R-mRNA)*,* either as a full-length mRNA (R-mRNA^{FL}) or as shorter nascent RNA sequences emerging during transcription as a platform for binding the first 30S subunit that initiates the pioneering round of translation (Figure 4.1D)[149]. Exploiting the high sequence conservation and comparable thermodynamics of the SD:anti-SD interactions between *Tte* and *Escherichia coli* (*Eco*; Figure 4.1E), we used *Eco* 30S subunits for comparing initiation complex (IC) formation efficiencies on the different-length of R-mRNAs. As expected, we found similar efficiencies for R-mRNA^{FL} and the RNA truncated 30 nucleotides downstream of the start codon

(R-mRNA^{+30}), but observed no IC formation on the RNA lacking both SD sequence and start codon (R-mRNA^{-11}; Appendix Figure 6.4). To avoid non-specific 30S binding downstream of the R-mRNA during the SiM-KARB assay and focus on the role of the 5'-UTR emerging first from transcription, we chose R-mRNA^{+30} as our nascent model mRNA transcript for studying 30S binding.

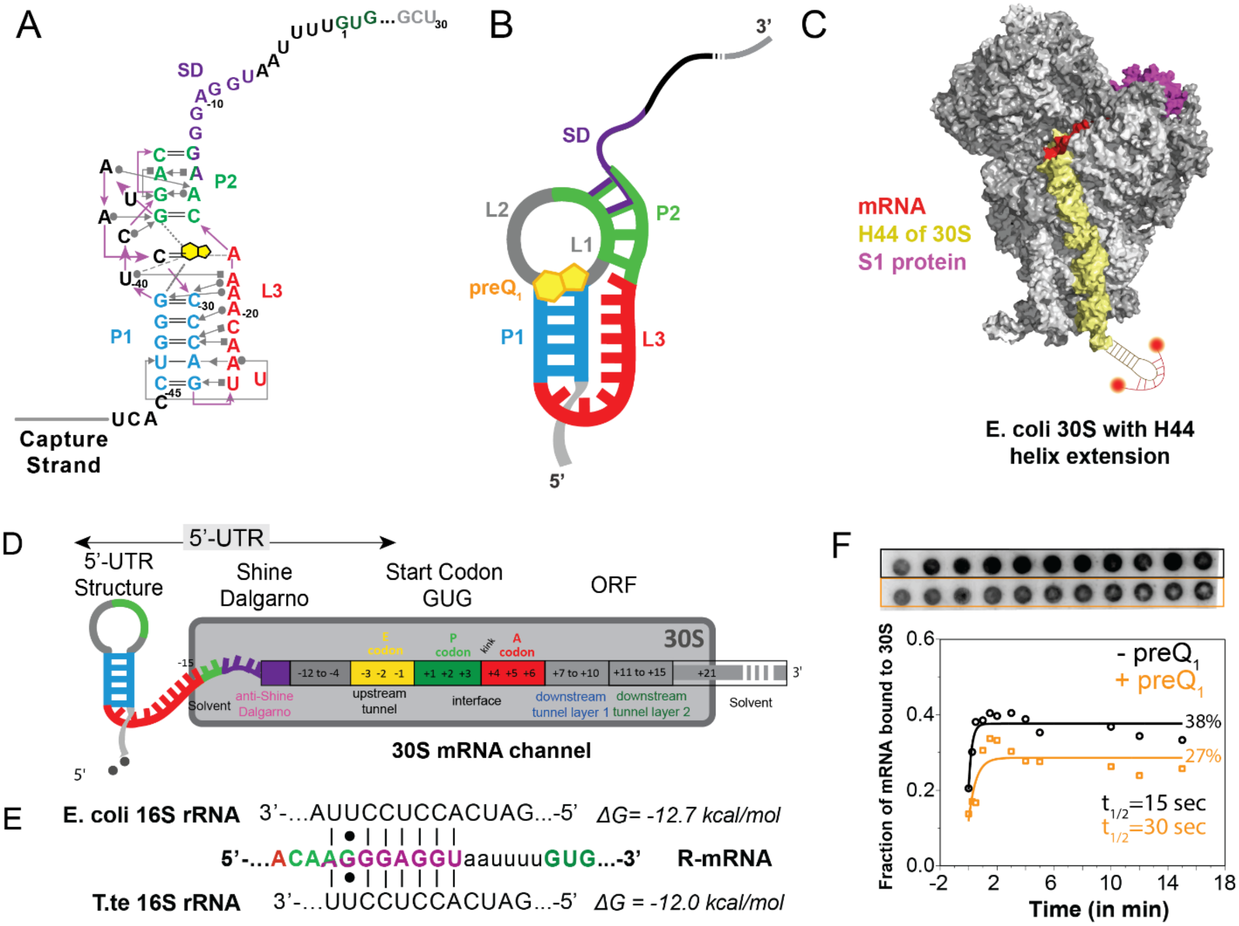


Figure 4.1 Schematic of a model mRNA with *Tte*-riboswitch at its 5'-UTR to study translation initiation.

(A) Structural map of the preQ₁ translational riboswitch from *Tte* displayed with Leontis–Westhof notations. The SD sequence (purple) partially overlaps the P2 stem nucleotides (green). (B) Schematic diagram of preQ₁ bound riboswitch containing mRNA (R-mRNA) from *T. tengcongensis (Tte)* (C) Schematic representation of the mutated ribosome, with a hairpin extension at the helix-44. The hairpin extension is hybridized with a dual labeled complementary oligo (in red). (D) Schematic diagram of *Tte* R-mRNA and its expected occupancy in the mRNA channel of the 30S during initiation, where +1 is the first nucleotide of the open reading frame. mRNA channel representation is adapted from[195]. (E) Base-pairing interactions between the Shine-

Dalgarno sequence of *Tte* mRNA transcript and rRNA of *E. coli* or *Tte* ribosomes. (F) Autoradiograph of filter binding membranes. Efficiency of 30S-mRNA complex formation as a function of time at no and 1 µM preQ1 concentration (top). Fraction of bound mRNA is plotted as a function of time (bottom).

To characterize the 30S binding efficiency of R-mRNA^{+30} in response to the addition of the preQ$_1$ ligand, we performed a radioactive filter-binding assay[196], wherein the fraction of 30S-bound over total ^{32}P-radiolabeled R-mRNA^{+30} is calculated from the blotted, double-filtered dots (Figure 4.1F). A time course of the normalized 30S-mRNA complex formation in the absence and presence of saturating (10 µM) preQ$_1$ revealed an ~2-fold adverse effect of preQ$_1$ on both pre-steady state kinetics and endpoint (Figure 4.1F). This level of modulation is very similar to earlier findings that the translation of R-mRNA^{FL} is decreased *in vitro* by ~40% at saturating preQ$_1$ concentration[149], suggesting that the ligand affects translation primarily at the investigated 30S binding step. The molecular mechanism of this gene regulation, however, is not revealed by such bulk assays.

4.3.2 Riboswitch ligand both antagonizes 30S recruitment and accelerates 30S dissociation

To directly evaluate 30S binding to R-mRNA^{+30} as the apparent gateway to initiating translation, we developed SiM-KARB, wherein an *Eco* 30S subunit with an extended 16S rRNA helix-44 was fluorescently tagged[197] by hybridization with a doubly (to mitigate photobleaching) Cy5-labeled DNA oligonucleotide (Figure 4.1C). Repeat binding and dissociation of an excess of this 30S subunit in solution (red) to surface-immobilized 3'-Cy3 labeled R-mRNA^{+30} (green) was monitored by total internal reflection fluorescence (TIRF) microscopy as transient diffraction-limited co-localization events in translation-active 50 mM Tris-Polymix Buffer[198] (Figure 4.2A,B; Methods for details). A mixture of short and long binding events was observed, suggesting (at least) two distinct types of 30S-mRNA interactions, with the presence of preQ$_1$ modulating the initial events of translation (Figure 4.2B). To obtain more quantitative, comparative insights, we pursued three levels of further analysis. First, individual time traces were fit with two-state Hidden Markov Models (Figure 4.2B), and these HMMs arranged into raster plots sorted by relative occupancy with 30S subunit, in either the absence or presence of preQ$_1$ (Figure 4.2C). Second, the HMMs were used to extract $\tau_{unbound}$ and τ_{bound} dwell times that were best fit globally with double-exponential functions[199,200], revealing each two rate constants: $\tau_{unbound}$ gave rise to $k_{on,fast}$ and $k_{on,slow}$, whereas τ_{bound} yielded $k_{off,fast}$ and $k_{off,slow}$. Of these, $k_{on,slow}$ and $k_{off,slow}$ represented ~30-60% and ~70-90%, respectively, of all cumulative events and were found to be most significantly impacted by preQ₁ over a range of ligand concentrations (Figure 4.2D, Table 4.1&2). Of note, while dual Cy5-labeling prolonged the observation of bound 30S subunits, we corrected all rate constants for photobleaching of both fluorophores (Appendix Figure 6.5). Global fitting of the ligand dependence with non-cooperative Hill-equations yielded a half-saturation point of ~48 nM (Figure 4.2D), in reasonable agreement with measurements of the preQ₁ ligand binding to the isolated riboswitch[43,189] and demonstrating that preQ₁ binding to R-mRNA^{+30} is fully saturated at 1 µM.

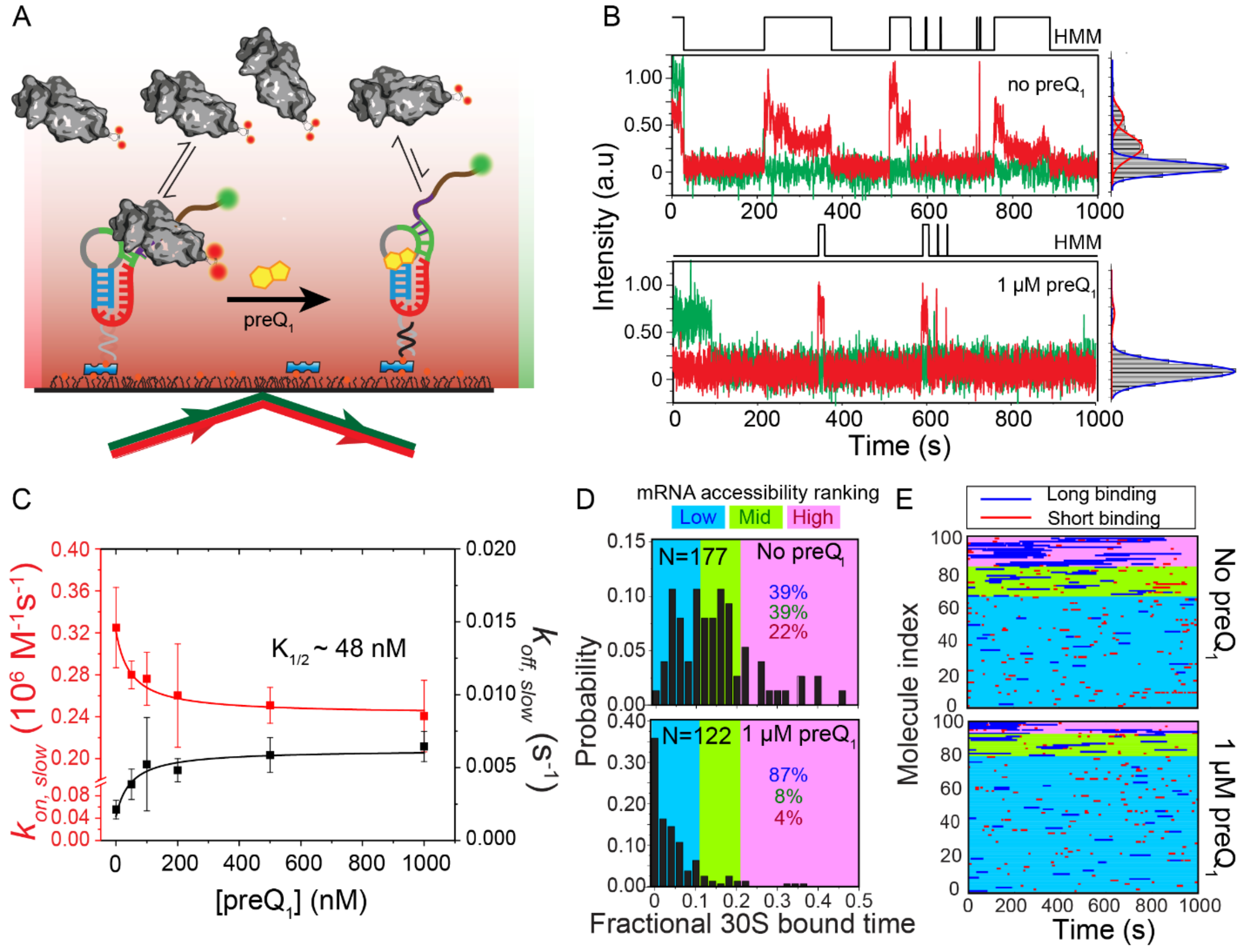


Figure 4.2 preQ₁-dependent 30S binding to R-mRNA at the single-molecule level.

(A) Schematic representation of the SiM-KARB assay. The Cy3 labeled R-mRNAs are immobilized to a slide surface. Repeated binding and dissociation of the 30S (dual-labeled with Cy5) is monitored through the co-localization of Cy3 and Cy5 fluorescence. (B) Representative R-mRNA^{+30} (green), and 30S (red)-binding fluorescence intensity versus time trajectory and corresponding intensity histogram for a single R-mRNA^{+30} molecule in the absence (top-right panel) and presence (bottom-right panel) of preQ$_1$. Hidden Markov Model idealization to a two-state model (black, top) are plotted as a function of time. (C) Association ($k_{on,slow}$, red) and dissociation ($k_{off,slow}$, black) rate constants for 30S interaction with R-mRNA^{+30} are determined from exponential fits of dwell times in the unbound and bound states, respectively, as a function of preQ$_1$ concentration. The corresponding $K_{1/2}$ values is determined by globally fitting $k_{on,slow}$ and $k_{off,slow}$ using Hill equation with Hill-coefficient as 1, respectively. (D) Histogram of the fractional bound-time for individual R-mRNA^{+30} in absence (top panel) and presence (bottom panel) of preQ$_1$. Based on the total fraction of bound time all 30S-bound R-mRNA^{+30} molecules were empirically ranked as low (<0.05 fractional bound time, cyan), mid (0.05 to 0.15 fractional bound-time, green), and high (>0.15fractional bound time, pink). N represents the totals number of molecules considered for each conditions (E) Raster plot of randomly selected 100 traces displaying the bursting behavior of 30S binding to the individual R-mRNA^{+30} based on the difference in accessibility for each molecule over time in the absence (top panel) and presence (bottom panel) of preQ$_1$. 30S binding to R-mRNA^{+30} is shown as bursts, represented in blue (short binding events or standby events) and red bar (long binding events or initiation-active events).

We first asked how 30S binding kinetics change at equilibrium when the riboswitch is pre-folded with varying concentrations of preQ$_1$ (between 0 nM to 1000 nM, Figure 2B, 2C). In the absence of preQ$_1$, we observe an association rate ($k_{on,slow}$) of $0.33 \pm 0.04 \times 10^6$ M^{-1}s^{-1} (relative amplitude = 89 ± 4%), and a dissociation rate ($k_{off,\ slow}$) of 0.002 ± 0.001 s^{-1} (58 ± 8%) (Figure 4.2C). The amplitude of rate constants showed that while $k_{on,slow}$ for 30S binding is close to homogenous, $k_{off,slow}$ is heterogeneous even after the global fitting. This apparent biphasic nature to $k_{off,slow}$ may be due to a broadly distributed molecular behavior resulting from the interaction between such macromolecules and corroborates with earlier observations[149]. In contrast, at saturating preQ$_1$ concentrations (1 μM), we observed a reduced $k_{on,slow}$ of $0.24 \pm 0.03 \times 10^6$ M^{-1}s^{-1} (96±8%) and an increased $k_{off,slow}$ of 0.006 ± 0.002 s^{-1} (41 ± 8%) (Figure 4.2C and Figure 4.3 and Table 4.1). A systematic study of the folded R-mRNA^{+30} with different preQ$_1$ concentrations revealed a ~35% decrease of $k_{on,slow}$ with increasing preQ$_1$. Such a decrease in $k_{on,slow}$ is expected because R-mRNA^{+30} is known to have higher accessibility to 30S binding due to the relatively open SD region in the absence of preQ$_1$. As preQ$_1$ concentration is increased, the riboswitch in the 5'-UTR folds to partially sequester the SD sequence. This is consistent with an accessibility competition model where the preQ$_1$ binding folds the 5'-UTR aptamer and competes with 30S binding to the mRNA, and preQ$_1$ binding to the aptamer precludes 30S from binding to the mRNA[22]. Furthermore, we noticed a three-fold increase in the dissociation rate ($k_{off,\ slow}$) with an increase of the preQ$_1$ concentration (Figure 4.2C and Figure 4.3 and Table 4.2). However, the increased $k_{off,slow}$ values do not correlate with a competition model as a higher concentration of preQ$_1$ should not affect the dissociation of the 30S subunit. This observation rather shows that preQ$_1$ not only competes for binding to the pseudoknot (evident from the $k_{on,slow}$) but also actively dislodges the 30S that is already bound to the R-mRNA^{+30}.

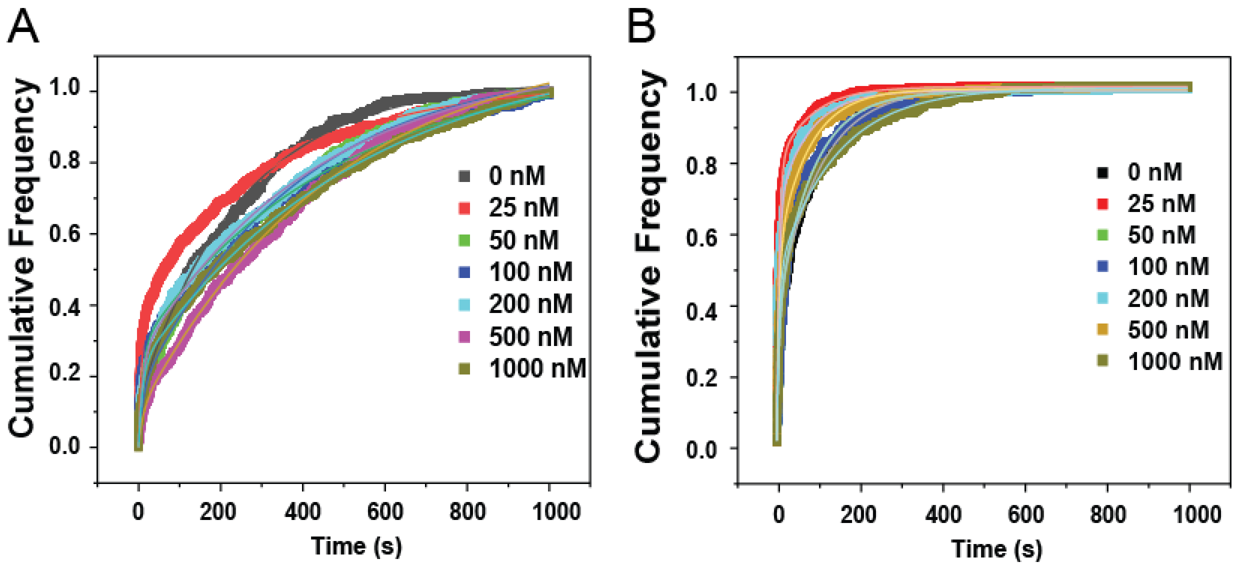


Figure 4.3 Plots of cumulative unbound (A) and bound (B) dwell times for the 30S binding at different concentrations of preQ$_1$

Table 4.1 List of two components unbound times ($t_{unbound}$) and on rates (k_{on}) for 30S binding to R-mRNA^{+30} at different concentration of preQ$_1$.

[PreQ$_1$] (nM)	*$t_{unbound,\ slow}$ (s)*	*±Δ $t_{unbound,slow}$ (s)*	*A_1*	*±ΔA_1*	*$k_{on,slow}$ (×10^6 M^{-1}s^{-1})*	*±Δ $k_{on,slow}$ (×10^6 M^{-1}s^{-1})*	*$k_{on,slow}$ (PB corrected) (×10^{6}M^{-1}s^{-1})*	*±Δ $k_{on,slow}$ (PB corrected) (×10^{6}M^{-1}s^{-1})*
0	*299.9*	*35.3*	*0.89*	*0.04*	*0.17*	*0.02*	*0.33*	*0.04*
50	*410.8*	*19.3*	*0.93*	*0.03*	*0.12*	*0.01*	*0.28*	*0.01*
100	*424.2*	*38.8*	*0.83*	*0.05*	*0.12*	*0.01*	*0.27*	*0.03*
200	*490.5*	*92.8*	*0.85*	*0.05*	*0.10*	*0.02*	*0.26*	*0.05*
500	*539.9*	*36.8*	*0.85*	*0.05*	*0.09*	*0.01*	*0.25*	*0.02*
1000	*607.5*	*86.1*	*0.96*	*0.08*	*0.08*	*0.01*	*0.24*	*0.03*

[PreQ$_1$] (nM)	*$t^{shared}_{unbound,fast}$ (s)*	*±Δ $t^{shared}_{unbound,fa}$ (s)*	*A_1*	*±ΔA_1*	*$k_{on,fast}$ (×10^6 M^{-1}s^{-1})*	*±Δ $k_{on,fast}$ (×10^6 M^{-1}s^{-1})*	*$k_{on,fast}$ (PB corrected) (×10^{6}M^{-1}s^{-1})*	*±Δ $k_{on,fast}$ (PB corrected) (×10^{6}M^{-1}s^{-1})*

0	*10.3*	*1.38*	*0.15*	*0.03*	*4.85*	*0.65*	*4.84*	*0.65*
50	*10.3*	*1.38*	*0.15*	*0.02*	*4.85*	*0.65*	*4.84*	*0.65*
100	*10.3*	*1.38*	*0.29*	*0.04*	*4.85*	*0.65*	*4.84*	*0.65*
200	*10.3*	*1.38*	*0.24*	*0.03*	*4.85*	*0.65*	*4.84*	*0.65*
500	*10.3*	*1.38*	*0.10*	*0.01*	*4.85*	*0.65*	*4.84*	*0.65*
1000	*10.3*	*1.38*	*0.22*	*0.01*	*4.85*	*0.65*	*4.84*	*0.65*

Table 4.2 List of all bound times (t_{bound}) and off rates (k_{off}) for 30S binding to R-mRNA^{+30} at different concentration of preQ$_1$.

[PreQ$_1$] (nM)	$t_{bound,slow}$ (s)	±Δ $t_{bound,slow}$ (s)	A_1	±ΔA_1	$k_{off,slow}$ (s^{-1})	±Δ $k_{off,slow}$ (s^{-1})	$k_{off,slow}$ (PB Corrected) (s^{-1})	±Δ$k_{off,slow}$ (PB Corrected) (s^{-1})
0	*137.2*	*41.3*	*0.58*	*0.08*	*0.007*	*0.002*	*0.002*	*0.001*
50	*110.8*	*29.9*	0.48	*0.07*	*0.009*	*0.002*	*0.004*	*0.001*
100	*96.2*	*58.9*	*0.41*	*0.03*	*0.012*	*0.002*	*0.005*	*0.003*
200	*100.1*	*16.3*	*0.48*	*0.06*	*0.010*	*0.002*	*0.005*	*0.001*
500	*90.7*	*18.5*	*0.49*	*0.01*	*0.007*	*0.001*	*0.006*	*0.001*
1000	*86.1*	*13.8*	*0.28*	*0.08*	*0.008*	*0.004*	*0.006*	*0.001*

[PreQ$_1$] (nM)	$t^{shared}_{bound,fast}$ (s)	±Δ $t^{shared}_{bound,fast}$ (s)	A_2	±ΔA_2	$k_{off,fast}$ (s^{-1})	±Δ $k_{off,fast}$ (s^{-1})	$k_{off,fast}$ (PB Corrected) (s^{-1})	±Δ$k_{off,fast}$ (PB Corrected) (s^{-1})
0	*7.2*	*1.1*	*0.36*	*0.03*	*0.138*	*0.021*	*0.133*	*0.020*
50	*7.2*	*1.1*	*0.51*	*0.07*	*0.138*	*0.021*	*0.133*	*0.020*
100	*7.2*	*1.1*	*0.58*	*0.08*	*0.138*	*0.021*	*0.133*	*0.020*
200	*7.2*	*1.1*	*0.51*	*0.07*	*0.138*	*0.021*	*0.133*	*0.020*
500	*7.2*	*1.1*	*0.38*	*0.10*	*0.138*	*0.021*	*0.133*	*0.020*
1000	*7.2*	*1.1*	*0.72*	*0.03*	*0.138*	*0.021*	*0.133*	*0.020*

To characterize how the accessibility of individual mRNAs change in response to preQ$_1$, we plotted accessibility histograms that bin the fraction of time 30S remains bound to each R-mRNA^{+30} molecules per experimental time window (Figure 4.2D). Based on three delineated average bound times in the pooled histograms from all data sets (Appendix Figure 6.6, Material and methods), the occupancy of individual R-mRNA^{+30} was categorized as Low (L, fractional 30S bound time <0.10), Mid (M, fractional 30S bound time between 0.10 and 0.20) or High (H, fractional 30S bound time >0.20) occupancy molecules. Without preQ$_1$, 39%, 39% and 22% of molecules show L, M, and H occupancy, respectively. In contrast, upon addition of 1 µM preQ$_1$, a large majority of the population (87%) exhibit L occupancy, whereas M (8%) and H (4%)

accessibility is significantly reduced (Figure 4.2D). These observations suggest that preQ$_1$ binding reduces the overall accessibility of the 30S to R-mRNA^{+30} and thereby hinder translation initiation (Figure 4.1A-C).

To understand the mechanism of this reduced accessibility, we looked at the individual 30S binding events to each R-mRNA^{+30}. The amplitudes of the double-exponential fits allowed us to quantify the percentage of long and short binding events in the context of preQ$_1$ binding. We found that the percentage of long and short binding events change from 60% and 36%, respectively in the absence of preQ$_1$ to 28% and 72% in the presence of 1µM preQ$_1$ (Table 4.2 & Figure 4.4). Thus, the relative fraction of long binding events decreases with the addition of preQ$_1$, while the fraction of short bindings increases. Previous reports have suggested that 30S subunit first forms a pre-initiation complex by nonspecifically binding to a "standby site". Once the 'proper' binding site is identified and the mRNA is accommodated into the 30S an "initiation-active" complex,which is formed for the next phase of translation[166,169,201]. To inquire whether the two different binding times (short and long binding) can be correlated to the standby and initiation-active binding, we plotted a raster plots (Figure 4.2E) that categorize individual binding events as short (red bar) and long (blue bar) binding based on a threshold bound time (Materials and methods). We further overlay the accessibility rank of the traces as L, M, or H, as described earlier. From these raster plots, we observed that the number of total binding events decreased by ~20% at 1 µM preQ$_1$. We further notice that the number of standby binding (decreased by ~20%) and the number of initiation-active binding (reduced by ~21%), both are affected in a relatively similar manner (Table 4.3). These data advocate for such a binding model consisting of random scanning of the mRNA followed by stable accommodation in the mRNA cleft, both of which are similarly impacted in response to preQ$_1$.

Overall, the SiM-KARB assay concludes that preQ$_1$ has a dual effect on the 30S binding. While on one hand, preQ$_1$ binding folds the aptamer to form a pseudoknot that impedes 30S binding, on the other hand, it partially sequesters the SD sequence to knock off bound 30S subunit from the R-mRNA^{+30}. As a combined effect, both standby and initiation-active binding events are impaired by the presence of preQ$_1$.

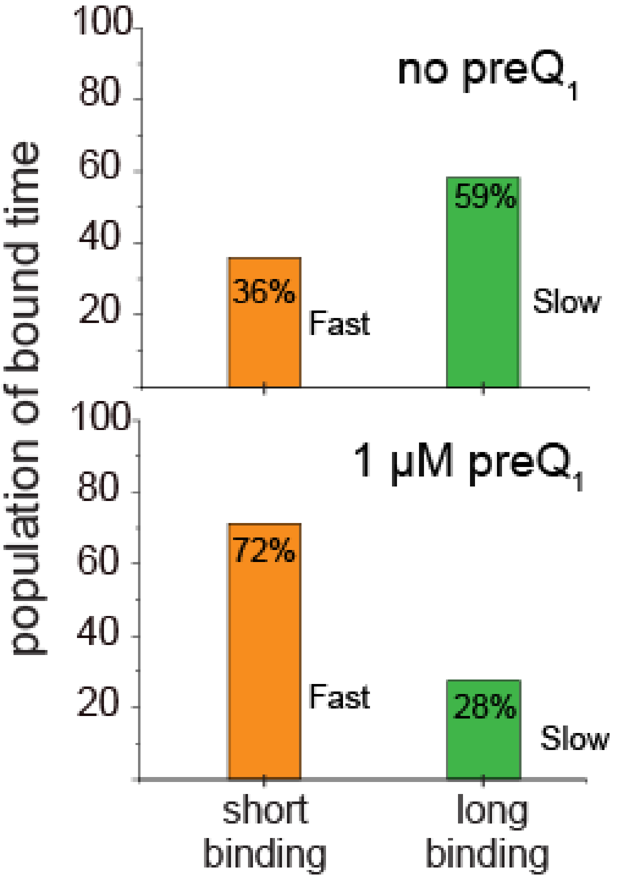


Figure 4.4 Percentage of long and short binding events for 30S binding to R-mRNA^{+30} in the absence and presence of preQ$_1$ determined from the biexponential fitting of the association (k_{on}) and dissociation (k_{off}) rates

Table 4.3 Counts of short (represented in red) and long (represented in blue) binding events from the raster plot of 100 molecules in the absence and presence of preQ$_1$ added to R-mRNA^{+30}

	Without PreQ$_1$				***With PreQ$_1$***			
	# of total molecules	*# of short binding*	*# of long binding*	*Total binding*	*# of total molecules*	*# of short binding*	*# of long binding*	*Total binding*
Low	*65*	*150*	*28*	*178*	*86*	*152*	*49*	*201*
Mid	*18*	*38*	*30*	*68*	*9*	*13*	*13*	*26*
High	*17*	*20*	*35*	*55*	*5*	*12*	*13*	*25*
Total	*100*	*208*	*93*	*301*	*100*	*176*	*75*	*251*

4.3.3 The 5'-UTR of R-mRNA dynamically adapts to changes in preQ$_1$ concentration

Next, we sought to understand how 30S binding to R-mRNA^{+30} changes in response to sudden variation in its environment (non-equilibrium), as may happen inside the bacterial cells due to environmental stress, salt, or small drug molecule concentration variation, etc[202]. For such case, we used non-equilibrium preQ$_1$ variation from no ligand to saturating ligand concentration to alter the structure of the 5'-UTR of the R-mRNA^{+30} in a "ligand-jump" experiment while simultaneously tracking 30S binding in a set of contiguous fluorescence traces (Figure 4.5A, Materials and methods)[149,203]. Similar to the equilibrium condition, we observed a reduction in 30S binding frequency (Figure 4.5A top and middle) and a decrease in the 30S bound time (Figure 4.5A, bottom). The decreased $k_{on,slow}$ and the increased $k_{off,slow}$ (Figure 4.6) reconfirms the observations from equilibrium observations hold for dynamic variation of preQ$_1$ in the micro-environment of R-mRNA^{+30}.

To understand the nature of 30S binding in such non-equilibrium conditions in detail, we plotted 30S fractional binding histograms using the traces from each conditions. In the absence of preQ$_1$, 45% of the R-mRNA^{+30} molecules were less accessible (L), while 31% and 24% of molecules showed medium (M) and high (H) accessibility, respectively (Figure 4.5B). By contrast, as saturating preQ$_1$ was added, the mRNAs predominantly exhibited reduced accessibility (82% L, 12% M and 6% L, Figure 4.5C). This shift from high accessibility to low accessibility with the addition of preQ$_1$ indicates that even in non-equilibrium conditions, resembling cellular environment *in vivo*, upon binding to preQ$_1$ single mRNA molecules repealed new 30S binding and expel bound 30S.

The raster plot allowed a direct comparison of the mRNA accessibility into the absence and presence of preQ$_1$ (Figure 4.5D). 48% of the population responded to the addition of preQ$_1$ with a reduction in accessibility from either high to medium or low accessibility, or from medium to low accessibility (Table 4.4). On the other hand, 46% remained in the same accessibility rank and did not respond to the addition of preQ$_1$. Surprisingly, ~6% population transitioned from a lower binding rank to a higher rank with the addition of preQ$_1$. This suggests that under non-equilibrium conditions, in some cases, preQ$_1$ may promote the refolding of an mRNA while allowing partial access of the SD sequence, as has been observed earlier[149]. While total 30S bound time was reduced for most molecules, some R-mRNAs (15%) became completely inaccessible to

30S binding. It was further observed that while the short standby site interactions reduced by ~30%, the longer initiation-active interactions reduced by ~60%, showing the profound effect of preQ$_1$ on initiation-active interactions (Table 4.5). These overall observations emphasized that the non-equilibrium structural changes at the 5'-UTR indeed occur quickly upon a change in the local environment and dictate the propensity of 30S binding.

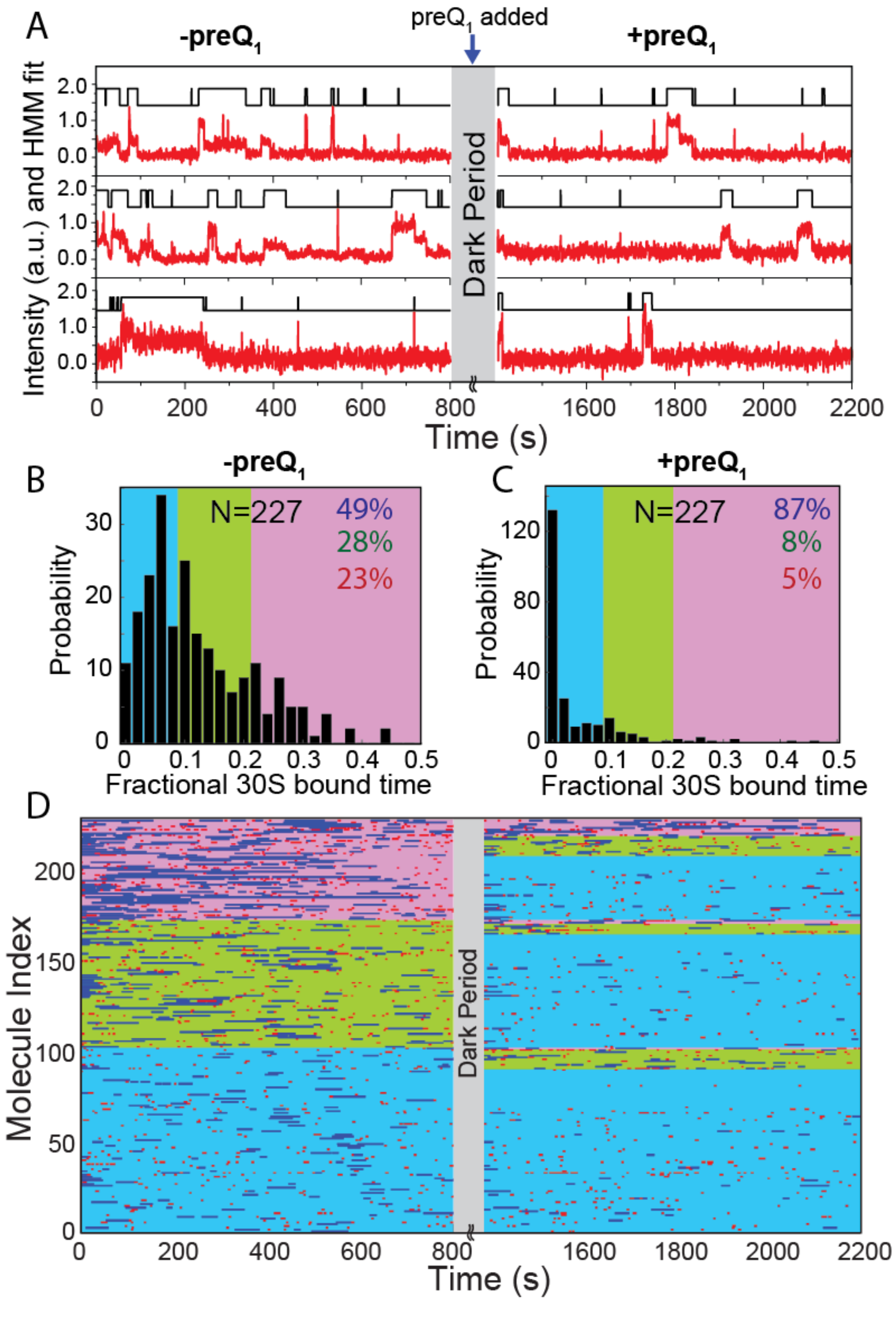


Figure 4.5 Single mRNA molecules undergo conformational switching upon the addition of preQ1 ligand.

(A) Exemplary single-molecule trajectories from non-equilibrium ligand-jump experiments composed of two segments. 30S probe binding to the same set of individual R-mRNA^{+30} is monitored first in the absence of preQ1 (left, -preQ1), then in the presence of 1 µM preQ1 (right, +preQ1). The gray axis break represents a 600 sec dark period between segments during which the buffer was exchanged. Histogram of the fraction of time individual R-mRNA^{+30} is bound by 30S before (B) and after (C) preQ$_1$ is introduced. Most molecules shift towards lower accessibility with the introduction of preQ$_1$. (D) Raster plot displaying the 30S binding behaviors based on different accessibility rankings for each segment of individual R-mRNA^{+30} molecules. 30S binding to R-mRNA^{+30} are shown as bursts, represented in blue (short binding events or standby events) and red bar (long binding events or initiation-active events).

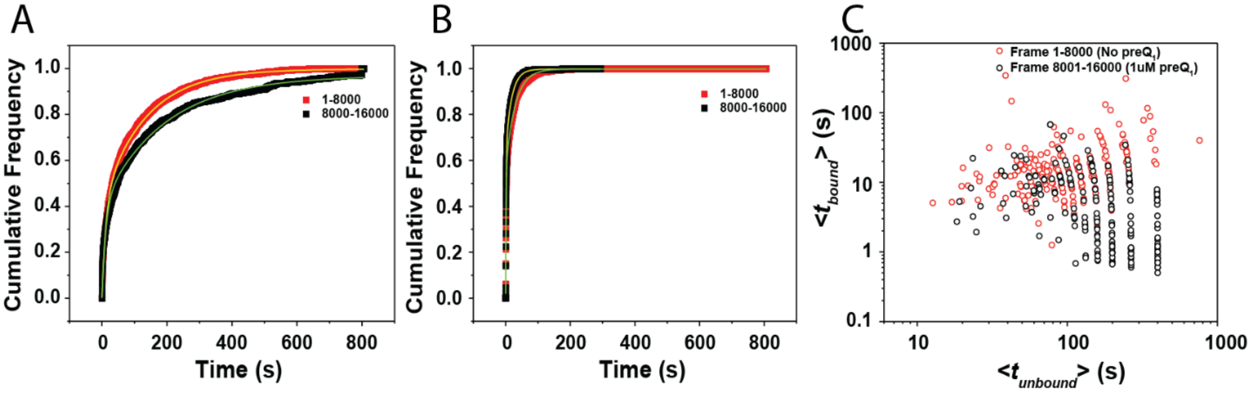


Figure 4.6 Plots of cumulative unbound and bound dwell times for the 30S binding without and with preQ$_1$

(a) Cumulative frequency plot for $t_{unbound}$ of 30S binding to R-mRNA^{+30} without preQ$_1$ (red) before dark period monitored for first 8000 frames and with 1 µM preQ$_1$ (black) after dark period monitored for next 8000 frames. 1 µM preQ$_1$ is added during the dark period that lasts for 5000 frames (1 frame = 0.1 sec) (b) Cumulative frequency plot for t_{bound} of 30S binding to R-mRNA^{+30} without PreQ$_1$ (red) and with 1 µM preQ$_1$ (black) (c) Scatter plot between average unbound time ($t_{unbound}$) vs average bound time (t_{bound}) obtained from the non-equilibrium ligand-jump experiment showing an increase in avg. $t_{unbound}$ and t_{bound}.

Table 4.4 Transition of accessibility with and without preQ$_1$ in the ligand jump experiment

Without PreQ$_1$ (Before ligand flow)		Transition	**With PreQ$_1$ (After ligand flow)**		
			Low	Mid	High
Low	102 (45%)		90	11	1
Mid	70 (31%)		62	6	2
High	55 (24%)		35	11	9
			187 (82%)	28 (12%)	12 (6%)
46% molecules stay in the same rank (Low→ Low, Mid→ Mid, High → High). *48% molecules have reduced accessibility rank (Mid→ Low, High → Low/Mid).* *6% molecules have increased accessibility rank (Low→Mid/ High, Mid→ High).*					
15% molecule transition to completely inaccessible 30S binding once preQ$_1$ is introduced.					

Table 4.5 Counts of short (represented in red) and long (represented in blue) binding events from the raster plot in the absence and the presence of 1µM preQ$_1$.

	Without PreQ$_1$ (before ligand flow)				With PreQ$_1$ (after ligand flow)			
	# of total molecules	# of short binding	# of long binding	Total binding	# of total molecules	# of short binding	# of long binding	Total binding
State1	45	151	62	213	82	211	41	252
State2	31	157	81	239	12	99	40	139
State3	24	197	111	308	5	45	27	73
Total	100	506	254	760	100	356	107	463

4.3.4 Appropriate positioning of the 5’-UTR structure provides hair-triggered precision for 30S binding

Next, we explored specific factors that control 5’-UTR accessibility, such as the strength of SD:anti-SD complementarity in the RBS and its proximity of the 5’-UTR influencing 30S binding. *In vivo,* this riboswitch and the SD sequence are intricately entangled in such a way that preQ$_1$ binding promotes P2 helix formation sequestering two nucleotides of the SD sequence[43]. To understand the role of the entangled aptamer and SD sequence, we strategically mutated the R-mRNA^{+30} to decouple these two functional parts. The first set of mutations allowed a systematic study of SD:anti-SD mediated 30S recognition, by varying lengths of the SD sequence complementarity from eight to no bases, while maintaining a four-base insert, thereby minimizing the influence of the aptamer (Figure 4.7A, I4S8 to I4S0; I = Insert length and S = SD length). In the second set of mutations, the aptamer and the SD sequence were systematically separated from each other by inserting one to six bases between them while always maintaining eight bases of SD sequence (Figure 4.8A, I1S8 to I6S8; I = Insert length and S = SD length). To eliminate any ambiguity in aptamer folding due to competition with the SD sequence, all mutants were designed such that the aptamer always retained all bases required for complete folding as well as the SD sequence always remained decoupled from the aptamer.

Using the SiM-KARB assay for mutants refolded under equilibrium conditions, we estimated a $k_{on,slow}$ value of $0.36 \pm 0.006 \times 10^6$ M^{-1}s^{-1} for I4S8,which decreased to $0.26 \pm 0.008 \times 10^6$ M^{-1}s^{-1} for I4S0 as the number of complementary base pairing between the SD and anti-SD (S8 to S0) decreased (Figure 4.7B, Appendix Table 6.2). This ~35% lowering in $k_{on,slow}$ with reducing SD:anti-SD complementarity confirmed that the availability of the SD sequence is an essential thermodynamic and kinetic factor for 30S recognition and binding to mRNA. Interestingly, we noticed the 30S still binds to the I4S0 construct even in the absence of SD sequences, which corroborates with earlier studies through other mechanisms like in leaderless mRNA[204,205]. We also observed that the $k_{on,slow}$ for I4S0 was close to the WT construct in the presence of preQ$_1$, providing the evidence that a similar level of blockage is observed in the absence of SD:anti-SD pairing (Figure 4.7B). As the SD sequence provides a direct recognition site for the 30S, the accessibility of the mRNA decreased as the number of base pairing were reduced. The fractional occupancy histograms showed that the mRNA accessibility progressively decreased from 15% high, 23% medium, and 62% low population for I4S8 to 5% high, 12% medium, and 83% low populations for I4S0 (Figure 4.7C). Furthermore, the raster plots for I4S8 and I4S0 showed that while the initiation-active (long) binding events were reduced by ~35%, the short standby bindings were reduced by ~45% (Figure 4.7D, Table 4.6). Overall, our results strongly suggest the role of SD:anti-SD pairing for initial docking of the 30S is crucial, and its elimination profoundly reduces overall 30S binding to mRNA and, as a result also impacting both the standby as well as the initiation-active binding.

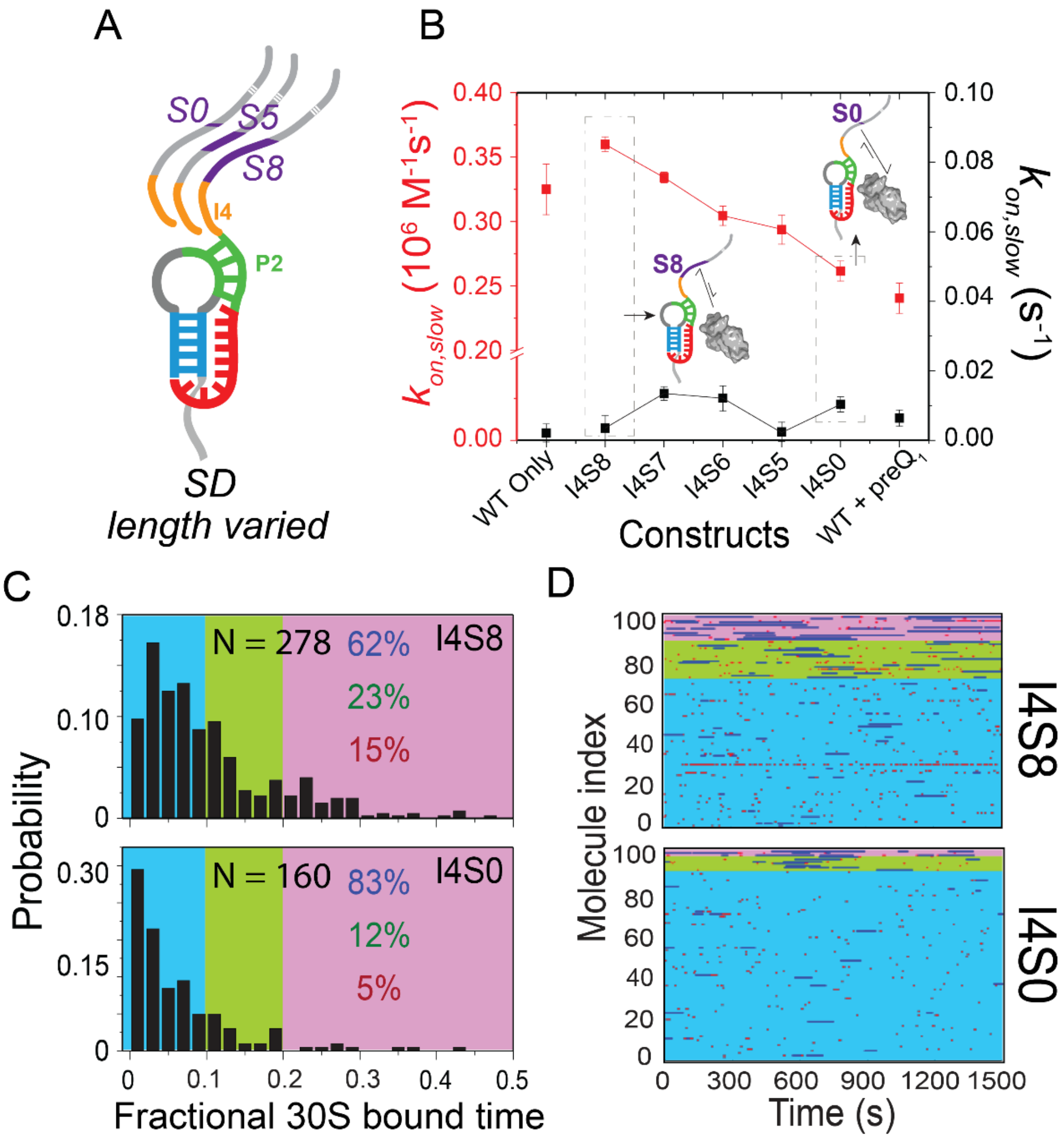


Figure 4.7 Mutation of the 5'-UTR structure by lowering SD sequence complementary (S) showed lower binding 30S binding.

(A) Schematic for varying the SD:anti-SD interaction length from full complementarity (S8) to no complementarity (S0) with a fixed number of bases (I4) between the aptamer and the SD sequence. (B) The association (k_{on}, red) and dissociation (k_{off}, black) rate constants for the 30S binding to the R-mRNA mutants with different SD:anti-SD complementarity. Data for the WT R-mRNA in the absence and presence of preQ$_1$ is shown for comparison with the mutant constructs. Binding kinetics of the 30S with the I4S8 and I4S0 mutants are represented in the schematics. (C) Histogram of the fractional bound-time for individual R-mRNA for I4S8 (top) and I4S0 (bottom). (D) Raster plot of randomly selected 100 traces displaying the bursting behavior of 30S binding to the individual mutant R-mRNA based on the difference in accessibility for each molecule over time in mutant construct with fully available SD (S8) complementarity (top panel) and no (S0) complementarity (bottom panel). 30S binding to R-mRNA is shown as bursts, represented in blue (short binding events or standby events) and red bar (long binding events or initiation-active events).

Table 4.6 Counts of short (represented in red) and long (represented in blue) binding events from the raster plot of 100 molecules for the two sets of mutants with fully available SD-aptamer complementarity (I4S8) to no SD-aptamer complementarity (I4S0)

	I4S8				**I4S0**			
	# of total molecules	# of short binding	# of long binding	Total binding	# of total molecules	# of short binding	# of long binding	Total binding
State1	73	306	17	323	89	226	16	242
State2	16	119	28	146	8	28	14	42
State3	11	65	24	89	4	14	14	28
Total	100	490	68	558	100	269	44	312

To understand the effect of the proximity of the aptamer in 5’-UTR, we studied the second set of mutants, I1S8 to I6S8 (Figure 4.8). We estimated a $k_{on,slow}$ value of $0.25 \pm 0.009 \times 10^6$ $M^{-1}s^{-1}$ for I1S8 which, very closely resembles that of the WT construct with saturating preQ$_1$ ($0.24 \pm 0.01 \times 10^6$ $M^{-1}s^{-1}$) (Figure 4.8B). When the fully folded aptamer is only one base away, even if all eight bases of the SD sequence are available for base pairing, it still effectively repeals 30S binding, resembling the WT construct folded with preQ$_1$. We hypothesize that this repeal of 30S must arise from steric hindrance between the aptamer and the 30S, which hinders 30S from accessing the SD sequence even though it is fully available for base pairing. To test the steric hindrance of the aptamer, we progressively moved the aptamer away from the SD sequence, expecting a recovery in the 30S binding propensity (mutant I1S8-I6S8). A similar $k_{on,slow}$ ($0.34 \pm 0.007 \times 10^6$ $M^{-1}s^{-1}$) compared to the WT construct without preQ$_1$ ($0.33 \pm 0.02 \times 10^6$ $M^{-1}s^{-1}$) was recovered when the aptamer is moved away by six insert length (I6S8, Appendix Table 6.4). Mutants imitating the intermittent distances (I2S8 and I4S8) showed a progressive impact on $k_{on,slow}$, suggesting that the proximity of the aptamer is very precisely regulated. The accessibility of the mRNA decreased when the aptamer is proximal to the SD sequence in I1S8 (9% high, 17% medium, and 74% low) and increased when the aptamer is more spaced out, as in the I6S8 construct (22% high, 30%

medium and 48% low) (Figure 4.8C). Furthermore, we found a ~165% increase in short standby binding events and ~36% decrease in long binding events of I6S8 compared to I1S8 (Figure 4.8D, Table 4.7). These observations showed that while the steric clash between the aptamer and the 30S influences initial 30S recognition, suitable positioning of the aptamer stabilizes the bound 30S. To eliminate the effect of aptamer and its influence on SD region due to preQ$_1$, one mutant from each set (I4S8 and I1S8) was used to verify that the preQ$_1$ has none/minimal effect on the kinetics of the mutants (Figure 4.9 and Appendix Table 6.6-6.9).

Observations from both sets of mutations suggested that while SD:anti-SD base pairing is essential for both initial recognition of the 30S as well as stable initiation-active binding, it is not the sole factor for mRNA accommodation[204]. Indeed, the steric hindrance between the 30S and the folded aptamer has a more profound effect on the initial standby binding. It is the interplay between the steric hindrance due to the proximity of the aptamer folding and the extent of SD sequence availability that have created a hair-triggered kinetic modulation of translation initiation.

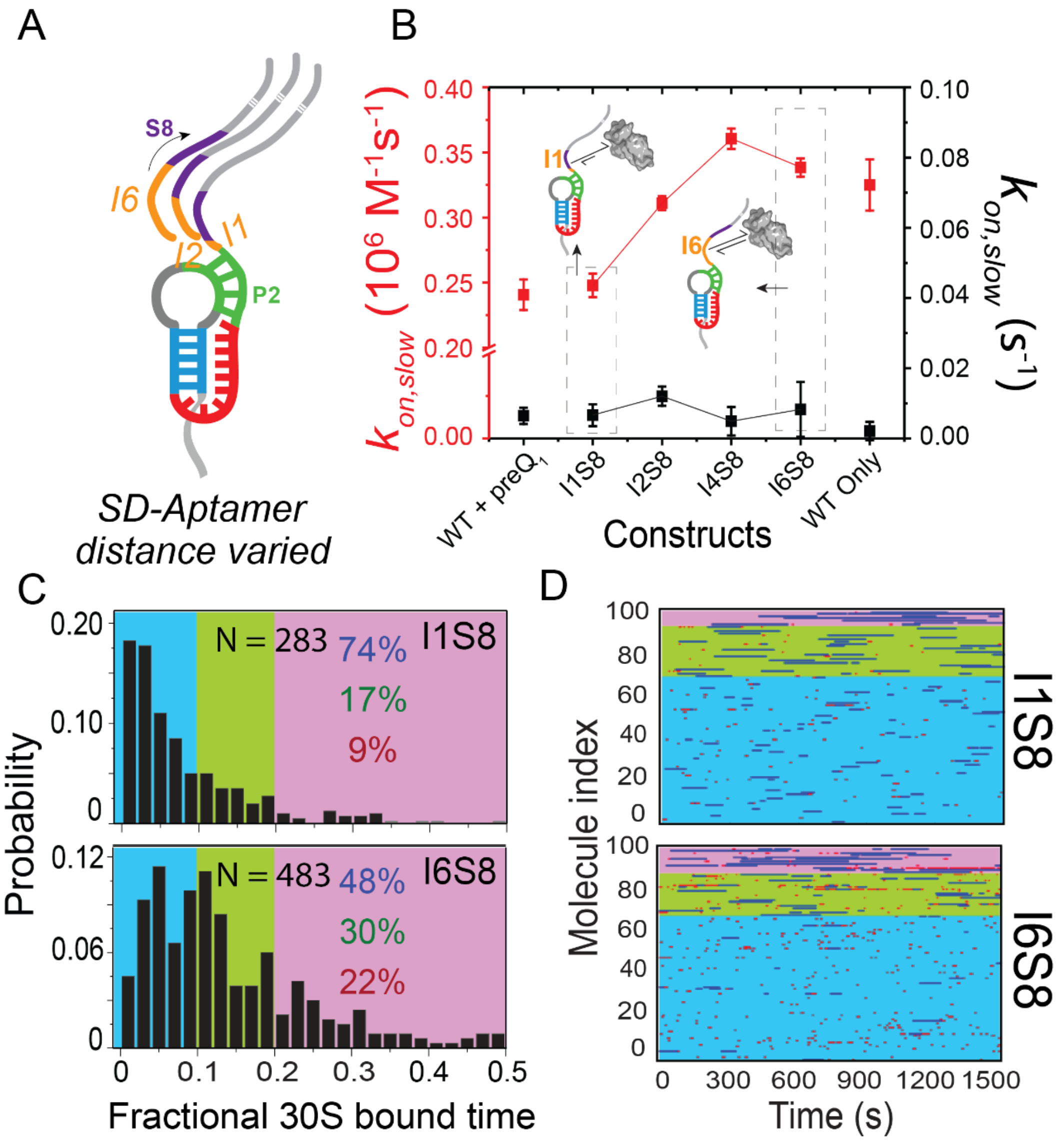


Figure 4.8 Mutation of the 5’-UTR structure by varying the SD-aptamer distance (I) from SD sequence (S) increased 30S accessibility to mRNA.

(A) Schematic for varying the SD-aptamer distance from insert length of one base (I1) to six bases (I6), while keeping SD sequence fully available (S8) for 30S binding. (B) The association ($k_{on,slow}$, red) and dissociation ($k_{off,slow}$, black) rate constants for the 30S binding to the R-mRNA mutants with different SD-aptamer distances. Data for the WT R-mRNA in the absence and presence of preQ$_1$ is shown for comparison with the mutant constructs. Binding kinetics of the 30S with the I1S8 and I6S8 mutants are represented in the schematics. (C) Histogram of the fractional bound-time for individual R-mRNA for I1S8 (top) and I6S8 (bottom). (D) Raster plots of randomly selected 100 traces displaying the bursting behavior of 30S binding to the individual mutant R-mRNA based on the difference in accessibility for each molecule over time in mutant construct with 1 base insert (I1) between aptamer and SD sequence (top panel) and with 6 base insert (I6) between aptamer and SD sequence (bottom panel). 30S binding to R-mRNA are shown as bursts, represented in blue (short binding events or standby events) and red bar (long binding events or initiation-active events).

Table 4.7 Counts of short (represented in red) and long (represented in blue) binding events from the raster plot of 100 molecules for the two sets of mutants with minimum SD-aptamer separation (I1S8) to maximum SD-aptamer separation (I6S8)

	I1S8				**I6S8**			
	# of total molecules	*# of short binding*	*# of long binding*	*Total binding*	*# of total molecules*	*# of short binding*	*# of long binding*	*Total binding*
State1	68	148	42	190	75	371	17	388
State2	22	47	34	81	15	110	25	134
State3	10	7	22	30	10	58	21	79
Total	100	203	98	301	100	538	63	602

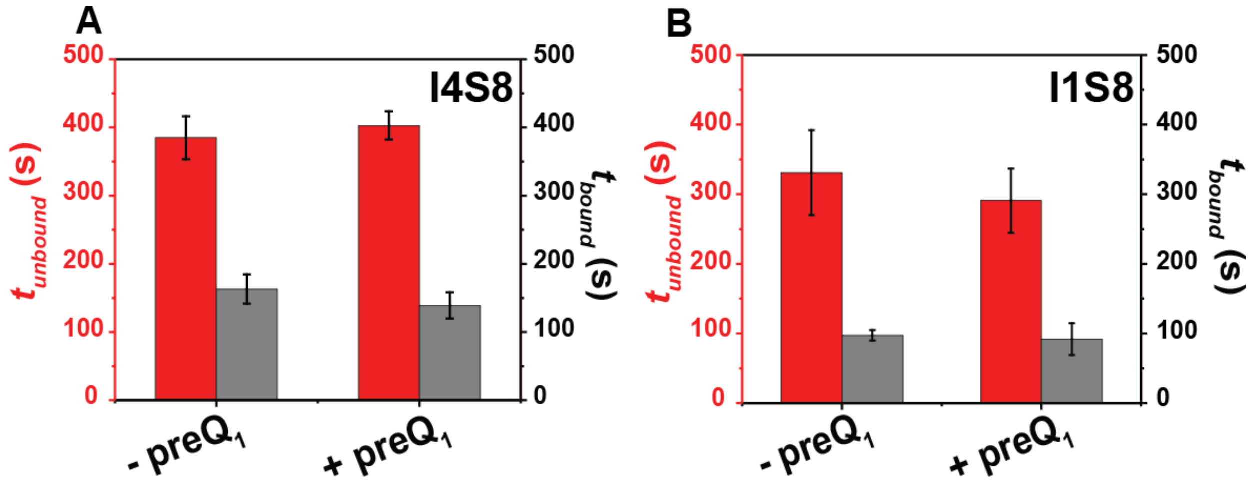


Figure 4.9 PreQ₁ binding did not affect the bound and unbound times for 30S binding and unbounding on mutant constructs I4S8 and I1S8

(A) Comparison of unbound ($t_{unbound}$, red) and bound times (t_{bound}, gray) for I4S8 in the absence and presence of preQ₁. (f) Comparison of unbound ($t_{unbound}$, red) and bound (t_{bound}, red) time for I1S8 in the absence and presence of preQ₁.

4.3.5 Initiation factors (IFs) and tRNA facilitating 30S binding but remains agnostic to 5'-UTR structures

Lastly, we explored how IFs and initiator fmet-tRNA^{fmet}, which are known to control translation initiation *in vivo*, impact 30S binding in the context of the 5'-UTR structure[206,207].

While each IF governs a specific function in the initiation process[208,209], we examined whether they aid in influencing 30S binding by interacting with the structured 5'-UTRs. We pre-mixed purified IF1, IF2, and IF3, fMet-tRNA^{fMet} simultaneously with the 30S in the tris-polymix buffer (Figure 4.10A). The mixture of IFs, tRNA, and 30S was introduced into the SiM-KARB assay and the binding kinetics were determined. We found that the mRNA accessibility in the presence of IFs and tRNA changed from 11% high, 21% medium, and 68% low accessibility without preQ$_1$ to 1% high, 4% medium, and 94% low accessibility with 1 µM preQ$_1$. This change in mRNA accessibility in response to preQ$_1$ is not drastically different from the data we obtained earlier in the absence of IFs and fmet-tRNA^{fmet} (Figure 4.10B).

The presence of IFs and fmet-tRNA^{fmet} enhanced the $k_{on,slow}$ to $0.39 \pm 0.01 \times 10^6$ M^{-1}s^{-1} and the addition of preQ$_1$ lowered the $k_{on,slow}$ to $0.27 \pm 0.04 \times 10^6$ M^{-1}s^{-1} while the $k_{off,slow}$ remained similar (Appendix Table 6.10). Our results revealed that IFs and tRNA^{fmet} increased the $k_{on,slow}$ both in the presence and absence of preQ$_1$. However, the relative effects prevail comparable to the binding of the 30S in their absence (Figure 4.10, Supplementary Table S13, Appendix Table 6.11). With these observations, we concluded that the effect of the 5'-UTR structure is analogous for the 30S binding regardless of the presence or absence of IFs and tRNAs.

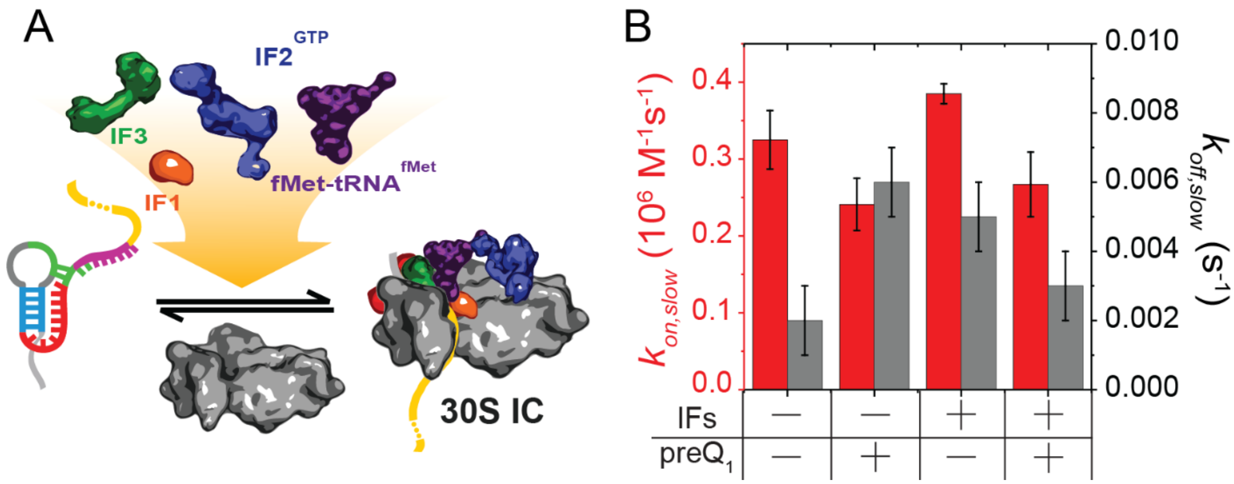


Figure 4.10 Addition of initiation factors (IFs) did not dramatically change the nature of 30S binding to the R-mRNA^{+30}

(A) Schematic of pre-initiation complex formation in the presence of initiation factors and the initiator tRNA. (B) Comparison of association ($k_{on,slow}$) and dissociation ($k_{off,slow}$) rates in the presence and absence of IFs with tRNA and/or preQ$_1$, showed a small effect on 30S binding.

4.4 Discussion

By examining the early stages of translation initiation, we demonstrated how a small but strategically placed 5'-UTR structure can influence gene expression. We demonstrated that ribosome binding to the mRNA is regulated by fine-tuned kinetic variations at the 5'-UTR. PreQ$_1$ binding to the aptamer at the 5'-UTR triggers a bifold mechanism to precisely tune the accessibility of 30S, repealing 30S binding via aptamer folding and releasing bound 30S via partial SD sequestration. This study highlights factors beyond start codon, such as the 5'-UTR structures, and the complementarity of SD:anti-SD base pairing can play a major role for such kinetic-tuning[62,210]. Using different mutants, we disentangled the 5'-UTR structure from the SD sequence and demonstrated that the proximity of the aptamer is the prime determinant of the initial recognition of the 30S while the extent of SD-mediated base pairing is essential for both initial and stable recruitment of the mRNA. The interplay between the steric hindrance due to proximity of the aptamer and the extent of SD availability, implement rapid standby site exploration followed by transient unfolding of the 5'-UTR respectively to provide precise hair-triggered kinetic modulation of translation efficiency. Furthermore, our work features the contrasting role of a small molecule preQ$_1$ that binds and stabilizes the riboswitch structure at the 5'-UTR.

In bacteria, mRNA starts to recruit the 30S to initiates translation as soon as it is being transcribed from RNA polymerase in an extremely kinetically controlled environment[211,212]. Highly structured RNA elements in the mRNA around the RBS are well known to facilitate this recruitment of 30S and precisely control translation efficiency in bacteria[166,169]. Structured 5'-UTR of mRNAs often act as roadblocks for mRNA accommodation in the small ribosomal subunit to form initiation complex (structures such as IRES, G-quadruplexes sometime act as recruiters)[213-215]. The classical interpretation of initiation on such structured mRNAs implies an equilibrium thermodynamic process where the 5'-UTR is remodeled into simpler low-energy structures by successive unfolding and refolding. This remodeled 5'-UTR then accommodates into the binding cleft of the 30S. However, meager changes in many 5'-UTR structures often liberate insufficient energy to surmount the energetic barrier of such mRNA remodeling. *In-vitro* translation of the full-length R-mRNA showed only a ~40% reduction in expression in the presence of saturating preQ$_1$, which is not a substantial variation as one would expect for a thermodynamically controlled binary model of riboswitch mediated gene regulation[185]. Recent transcriptome-wide studies have reformed such an equilibrium model and shown that a non-equilibrium kinetic competition between mRNA unfolding and 30S dissociation governs translation efficiency[166]. Such kinetic competition further manifests how motifs such as the riboswitches modulate initiation, despite minor variations in equilibrium stability. The small kinetic traps that prevent 30S binding, catalyze gene expression modulation by creating a deficit in the number of free ribosomes as they initiate on other mRNAs in the cell.

Based on the kinetic parameters measured using SiM-KARB assay, we predict a free energy model for the non-equilibrium unfolding of the riboswitch at 5'-UTR (Figure 4.11). We noticed that the initial binding energy is very similar both in the presence and absence of preQ$_1$ (-2.1 kcal/mol). The successive mRNA accommodation into the 30S binding cleft is a spontaneous process and releases (-) 0.25 kcal/mole excess energy in the absence of preQ$_1$. However, with preQ$_1$, the 5'-UTR is more stably folded and requires (+) 0.4 kcal/mol excess energy to unfold. Thus, the energetic penalty for initiation-active complex formation is ~0.8 kcal/mol higher when preQ$_1$ is present. This energetic penalty is even smaller than the energetic cost of a couple of hydrogen bond formation due to preQ$_1$ binding to the aptamer[190]. Thus, these observations fit well with our model, where the 30S binding is governed by the non-equilibrium/dynamic unfolding of the riboswitch and does not require complete unwinding of the 5'-UTR structure.

Such modulation is far from unique to the class-I preQ$_1$ riboswitch. In a recent study with Mn^{2+} sensing riboswitch, a single metal ion at the local binding core can alter the global structural change resulting in translation regulation[82]. It is perhaps more interesting that an uncharged ligand, as small as preQ$_1$ binding to a small riboswitch aptamer, can transduce such a global effect of fine-tuned translation regulation, triggering series of kinetic changes that influences the 30S binding and therefore prompts a tug of war against mRNA accommodation into the 30S mRNA cleft.

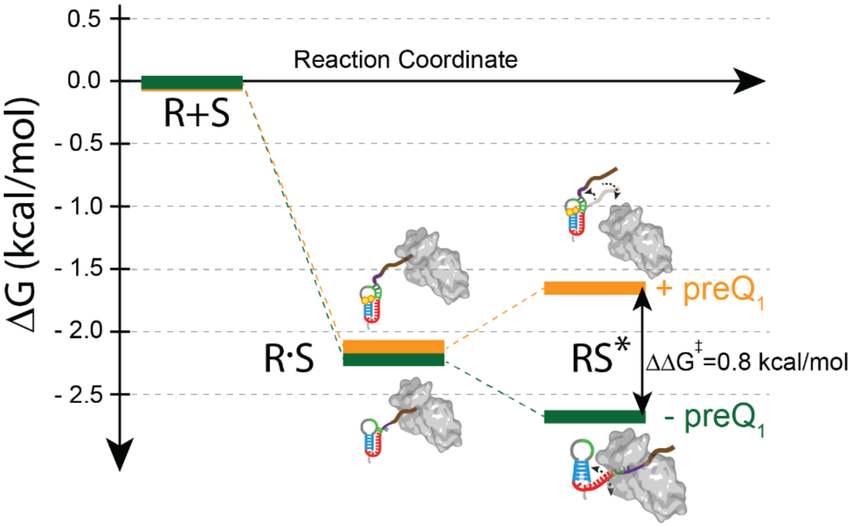


Figure 4.11 Free energy diagram represents the energy levels calculated from the rate constant values in the absence and presence of preQ$_1$ in the riboswitch at the 5’-UTR.

The 30S-mRNA binding model shows an energetic penalty of ~0.8 kcal/mol for the formation of initiation-active complex that is achieved through the dynamic unfolding of the 5’UTR leading to complete accommodation of the mRNA into the 30S.

Placing our observations in context with the prevailing knowledge, we propose a model for the binding of mRNAs with structured 5’-UTR to the 30S mRNA binding cleft (Figure 4.12). The 30S essentially waits for partial melting of the 5’-UTR for accessibility to the RBS through standby site interaction. These interactions are transient and highly dependent on ligand binding. In the absence of ligand, S1 promotes dynamic unfolding of the 5’-UTR and thus entropically favors full accommodation of the mRNA into the binding cleft of the 30S, thus transitioning to an initiation-active binding mode that subsequently converts into an active translation complex with the further addition of initiation factors (IFs) and initiator tRNA. However, in the presence of preQ$_1$, the aptamer is tightly folded, thus impedes initial 30S binding by allowing limited time for the dynamic unfolding of the aptamer. Moreover, preQ$_1$ actively promotes folding the 5'-UTR of the mRNA and thus hinders initiation-active binding by dislodging bound 30S.

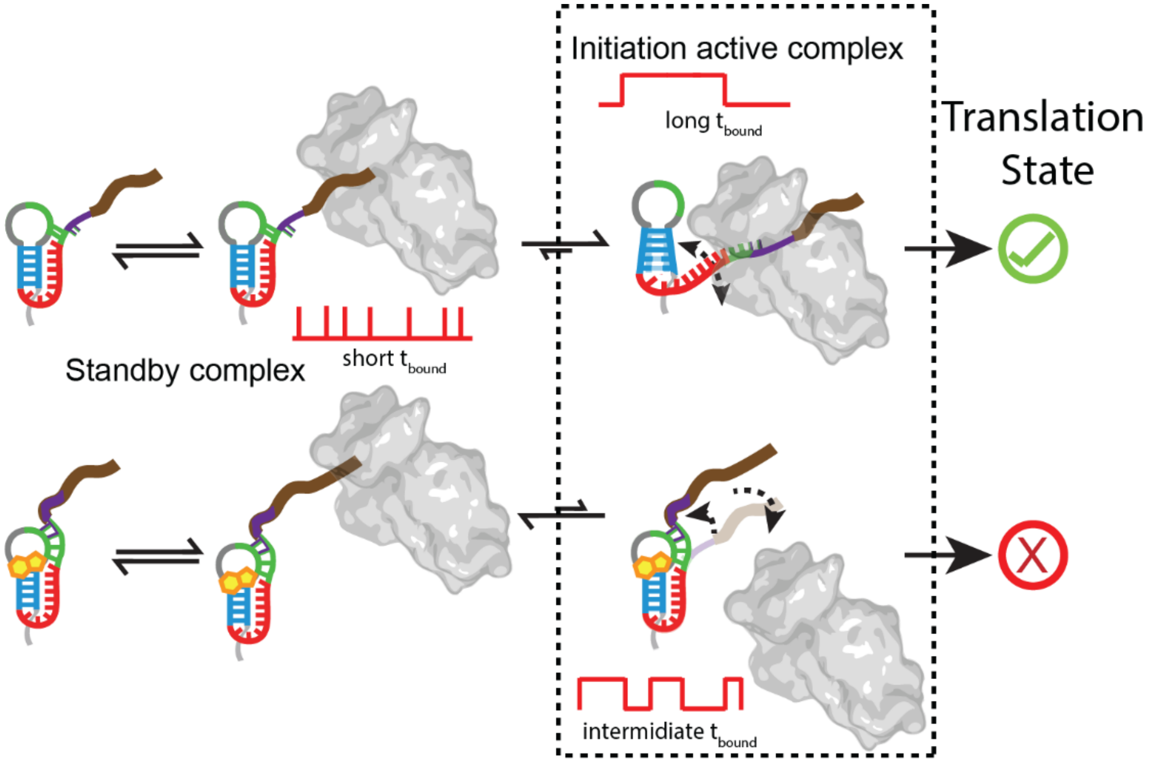


Figure 4.12 Model for 30S binding to the SD region of R-mRNA regulated by a preQ$_1$ riboswitch at 5'-UTR.

Independent of the preQ$_1$, the 30S dynamically interacts with the R-mRNA looking for the RBS (SD sequence) to facilitate translation initiation on the nascent R-mRNA. Short transient 30S binding events are categorized as standby binding, where 30S has not found the right RBS or the additional factors (right orientation around RBS, IFs, fmet-tRNA^{fmet}, etc.) required to initiate translation. In the absence of preQ$_1$, the 5'-UTR is in less dynamic interaction with the SD sequence, leading to an unsequestred SD region that when finds the 30S with complete accommodation of R-mRNA through correct SD:anti-SD interaction can lead to stable 30S binding, characterized as initiation-active binding events. However, in the presence of preQ$_1$, the 5'-UTR partially sequesters the SD sequence leading to partial accommodation of R-mRNA into the 30S, which may either prevent new 30S to bind or disrupt bound 30S to R-mRNA and result in unsuccessful translation.

In the bacterial cell, translation efficiency plays a significant role in determining the fate of transcribed mRNA. As the 30S initiation complexes form, commitment to translation increases while the vulnerability to mRNA decay decreases, as mediated by the bacterial degradosome, Hfq, and Rho[216-218]. Thus, strategically evolved kinetic modulation of initiation complex formation is the key regulator of global translation activity, either to repel the 30S subunit directly or in coordination with other forces in the cell. Our work involving a structured 5'-UTR to study 30S accessibility during the initial stage of translation initiation allows us to construct a comprehensive model of translation regulation in bacteria by measuring kinetic features of 30S-mRNA interaction. These results further pave the way to study more complex biological systems such as co-transcriptional translation by the expressome assembly or mRNA degradation by the degradosome.

4.5 Methods and Materials

4.5.1 Ribosome preparation

Mutant *E. coli* pKK3535 plasmid strain with an extension at the helix-44 of the 16s rRNA was obtained from Prof. Joseph Puglisi's laboratory. This extension allows labeling of the ribosome using a fluorescently labeled DNA oligonucleotide complementary to the extended portion of the helix-44 without affecting the ribosome's functionality[197]. The helix extension presence was further verified by sequencing the plasmid. To confirm that the strain contains only mutated ribosome where the 16S ribosomal RNA contains the extension, total RNA was extracted from a small culture of cells following previously published protocol. The extracted RNA was then verified by reverse transcription followed by DNA-sanger sequencing. Single salt-washed ribosomes were prepared using a previously described protocol with several modifications[197]. Briefly, the pKK3535 strain, containing mutated ribosome was grown in Luria-Bertani (LB) medium at 37 °C to an OD_{600} of 0.8-1 starting from an overnight culture. The cells were then cooled at 4 °C for 45 mins and pelleted at 5,000 rpm for 15 mins. All subsequent steps were performed on ice or at 4 °C. The cell pellet was resuspended in buffer A (20mM Tris-HCl (pH 7.05 at 25 °C), 100 mM NH_4Cl, 10 mM $MgCl_2$, 0.5 mM EDTA and 6 mM β-mercaptoethanol), and the cells were lysed in a single pass using a M-110L Microfluidizer processor (Microfluidics). The lysate was cleared by centrifugation at 16,000 rpm in a JA-20 rotor. The clarified lysate was pelleted over a 35-ml sucrose cushion (1.1 M sucrose, 20 mM Tris-HCl (pH 7.05 at 25 °C), 500 mM NH_4Cl, 10 mM $MgCl_2$ and 0.5 mM EDTA) in a Beckman Ti-45 rotor for 16-20 hours overnight at 37,000 rpm. The pellet was washed twice with 1ml of buffer B (20 mM Tris-HCl (pH 7.05 at 25 °C), 500 mM NH_4Cl, 10 mM $MgCl_2$ and 0.5 mM EDTA), resuspended in 6 ml of buffer B by gentle stirring. The one salt-washed 70S ribosomes were then dialyzed against low magnesium buffer E (50 mM Tris-HCl (pH 7.05 at 25 °C), 150 mM NH_4Cl, 1 mM $MgCl_2$ and 6 mM b-mercaptoethanol) three times. The low Mg^{2+} ions facilitate the two subunits to dissociate from each other and remain as individual subunits in the solution. Next, 100 A_{260} unit of the dissociated ribosome is loaded onto a 36 ml of previously prepared sucrose gradient (to prepare the gradient, buffer E + 20% sucrose is frozen at -80 °C and then gently thawed at room temperature without mixing). The gradients were then loaded onto a swinging bucket (SW-28) rotor and centrifuged at 20,000 rpm for 18 hours. The gradients were then fractionated using a Brandel gradient fractionator coupled with a UV-detection monitor. Appropriate fractions were pooled together as pure 30S and pure 50S fractions. The 30S and 50S fractions were then palleted separately for 12 hours at 66,000 rpm in a Beckman Ti-70 rotor. Pelleted subunits were resuspended in storage buffer (50 mM Tris-HCl (pH 7.5 at 25 °C), 70 mM NH_4Cl, 30 mM KCl, 7 mM $MgCl_2$ and 6 mM b-mercaptoethanol) and flash-frozen with liquid nitrogen and stored at -80 °C.

4.5.2 Cloning of sequences encoding the R-mRNAs and different mutants

The complete mRNA transcript, including the TTE_RS07450 and TTE_RS07445 (TTE1564 and TTE1563, respectively) ORFs, and its 30 UTRs as predicted from the FindTerm algorithm (Soft-Berry), was amplified using PCR from *T. tengcongensis* genomic DNA, which was purchased from the NITE Biological Resource Center. The amplified region was cloned into the pUC19 plasmid between the BamHI and HindIII sites with an engineered upstream T7 promoter (pUC19_*Tte*). Different lengths of DNA were prepared by PCR amplification of the desired parts of the DNA. DpnI enzyme, which cleaves methylated DNA, is used to digest parent plasmid. 2´-O-methylation (2´-OMe) modification in the first two bases of the reverse primers ensures that during transcription in the next step, RNA polymerase dissociates without adding additional bases at the 3' end of the RNA. Transcription reactions were performed in the presence of 120mM HEPES-KOH (pH 7.5 at 25 °C), 30 mM $MgCl_2$, 2 mM spermidine, 40 mM dithiothreitol (DTT), 30 mM NTPs, 0.01% (w/v) Triton X-100, 400 nM PCR amplified DNA, 0.01 U/ml pyrophosphatase and 0.2 mg/ml T7 RNA polymerase in a total volume of 150 µl. Transcription reactions were incubated at 37 °C for ~18 hours. mRNA was purified by denaturing, 7 M urea, PAGE, detected using brief 254 nm ultraviolet radiation and gel-eluted overnight. mRNAs were ethanol-precipitated and resuspended in water.

Different R-mRNA mutations were generated from the original pUC19_Tte plasmid. PCR based site-directed mutagenesis was performed with primers (Section 4.5.11 for DNA primer sequences) designed to span the mutated bases. RNAs were generated from the plasmids in the same way as described above.

4.5.3 RNA filter-binding assay

The principle for this assay is based on the double-filter method described before, in which the binding of radiolabeled nucleic acids by proteins or other macromolecules is assessed by filter binding reactions through a pair of stacked membranes and measuring the amount of radioactivity retained in each[196,219,220].

Radiolabeled R-mRNA^{+30} was prepared in two steps. In the first step, 5 μM RNA was dephosphorylated using Antarctic phosphatase enzyme (New England Biolab). The reaction was incubated at 37 °C for 30 mins, followed by heat deactivation for 2 mins at 80 °C. In the second step, dephosphorylated RNA was phosphorylated using ^{32}P-labeled ɣ-ATP and T4 Poly Nucleotide Kinase (NEB). The reaction was incubated at 37 °C for 30 mins. The resulting radiolabeled RNA was purified by a spin column.

2X salt-washed 30S subunits used for all assays were activated by incubation at 37 °C in the Tris polymix buffer for 5 mins immediately prior to use. For each reaction, 3 μL of 3 mM R-mRNA^{+30} was re-folded in the absence and presence of 1 μM preQ$_1$ by heating to 70 °C for 2 mins, followed by slow cooling to room temperature for 15 mins. 30S subunits in Tris Polymix buffer was added to the RNA with and without preQ$_1$ at different time interval into the filtration column. Membranes were pre-wet in binding buffer for 30 mins. A membrane stack was constructed by stacking (from top to bottom): a reinforced nitrocellulose membrane (Optitran BA-S 85, Whatman #10-439-191), a Whatman 1MM filter paper, a positively charged nylon membrane (BrightStar-Plus, Ambion), and a second Whatman 1MM filter paper. The membrane stack was clamped inside of 96-well dot-blot manifold (Mini-fold, Schleicher & Schuell) and were then washed with cold binding buffer (100 μL per well) and dried by applying vacuum.

The 30S and RNA binding reactions were pipetted into the wells of the manifold, drawn through the membranes under vacuum, and then washed 100 μL cold buffer. Vacuum was applied for 2-3 mins until the membranes appeared dry and then membranes were wrapped in saran wrap and imaged using a storage phosphor screen and Typhoon 9410 Variable Mode Imager (GE Healthcare Life Sciences) and quantified using ImageQuant v 5.2 (Molecular Dynamics).

Radiolabeled RNA that is successfully incorporated into the 30S-RNA complex is preferentially retained in the nitrocellulose filter, while unbound tRNA is trapped in the positively charged nylon filter. The fraction of bound and unbound 30S were calculated as described previously[220].

4.5.4 Single-molecule assay

A prism-type total internal reflection fluorescence setup built around an Olympus-IX83 microscope, equipped with 60× 1.20 N.A. water objective and four sCMOS cameras (Hamamatsu, Flash-4 V3) and four different wavelength laser lines was used to perform the ribosome binding experiments (Only two laser lines and two cameras are used). Flow cell sample channels were prepared on surface passivating quartz microscope slides though coated with a mixture of 90% methoxy Poly-Ethelyn Glycol succinimidyl valeric acid (m-PEG SVA) and 10% biotin-PEG succinimidyl valeric acid (biotin-PEG SVA) using previously established protocols[70,221]. For surface immobilization of R-mRNA, the sample chamber was treated with 0.2 mg/mL streptavidin to bind to the biotin from the PEG. Solution containing ~50 pM previously annealed biotinylated capture strand-RNA complex was introduced to the chamber to sparsely coat the PEG surface with streptavidin. Excess non-immobilized RNAs were then washed with 200 μL wash buffer (10 mM Tris Base, pH 7.5 @ 25 °C). Steady-state SiM-KARB and ribosome binding measurements were performed by first forming the riboswitches by incubating the surface-immobilized RNA constructs with at a given concentration of $preQ_1$ in Tris-polymix buffer (50 mM Tris-OAc (pH 7.5 @ 25 °C), 100 mM KCl, 5 mM NH_4OAc, 0.5 mM $Ca(OAc)_2$, 5 mM $Mg(OAc)_2$, 6 mM β-mercaptoethanol, 0.5 mM EDTA, 5 mM putrescine, and 1 mM spermidine) for 15 mins. 20 nM concentration of dual Cy-5 labeled ribosome solution with same concentration of $preQ_1$ was then added to the chamber in an imaging solution (1X Tris-polymix buffer, 5 mM protocatechuic acid and 50 nM protocatechuate-3,4-dioxygenase, 2 mM Trolox). The dual labeling ensured long binding events of 30S to R-mRNA despite photobleaching of individual Cy5 molecule. An integration time of 100 ms was used unless otherwise specified for the experiments. A combination of continuous and shuttered illumination was used to capture slow and fast dynamic events, respectively. Shuttered illumination was specifically used for non-equilibrium ligand-jump experiments. Ten thousand frame movies at the rate of 100 ms per frame (150 ms per frame for mutants due to long observed binding times of 30S to mutants) were recorded for each condition with continuous 532 nm and 639 nm laser sources for the whole duration.

4.5.5 Ligand-jump experiment

Surface immobilized R-mRNA^{+30} molecules in the absence of preQ$_1$ were first introduced to buffer containing 20 nM of Cy5 labeled 30S and monitored for eight hundred seconds, followed by six hundred seconds dark time (shown in gray, Figure 4.5A). During the dark period, a fresh buffer solution containing 1 µM preQ$_1$ and same concentration of Cy5-30S as before was flown into the channel. The dark period allowed time for the homogeneous exchange of buffer and OSS to reduce photobleaching probability without disturbing the field of view. After the dark period, molecules in the same field of view were tracked in real-time for another eight hundred seconds in the presence of the preQ$_1$ without altering the preexisting concentration of buffer and 30S in the solution.

4.5.6 Analysis of single-molecule data

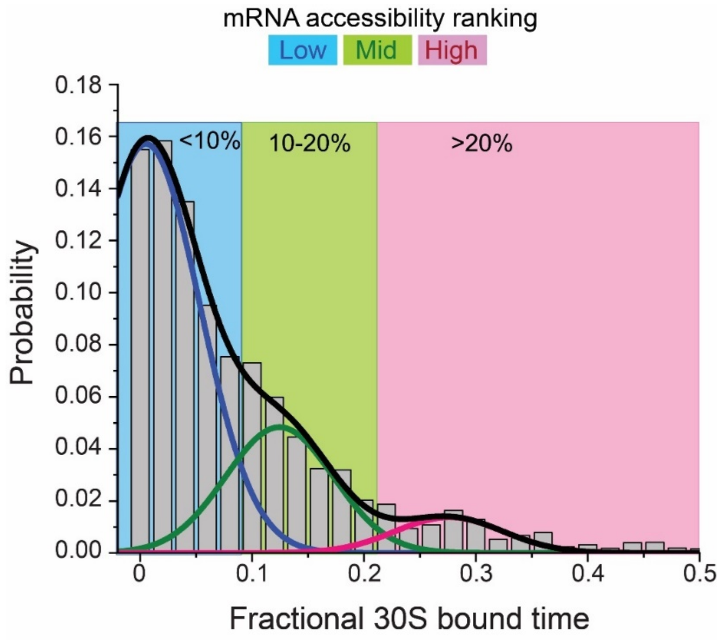
Single-molecule time traces were generated by a custom-written MATLAB code. Furthermore, custom analysis programs in MATLAB were used to extract statistical data from individual molecules, and finally Origin Pro was used to plot the data.

4.5.7 Global fitting

All the data for association and dissociation rates for each condition were fitted together globally in Origin pro software to get higher accuracy of double-exponential fitting[222,223]. To reduce the number of independent parameters associated with double-exponential fitting for both association and dissociation rates as well as considering the heterogeneity of the 30S binding to the R-mRNA^{+30} for each condition, the shorter binding time of the double-exponential fitting was shared across all the conditions of preQ$_1$ and effect of mutations. The global fitting of rates separately yielded two components for each of association ($k_{on,slow}$ and $k_{on,fast}^{shared}$) and dissociation ($k_{off,slow}$ and $k_{off,fast}^{shared}$) rates, one of which was variable and the other was shared over all the conditions. While the shared component of the rate constant (k^{shared}) remained fixed for all conditions, the variable component was observed to be profoundly responsive to the influence of preQ$_1$. The variable rate constant components and their relative contributions were used to compare different conditions to determine the role of preQ$_1$ and strategic mutations on 30S ribosome binding.

4.5.8 Distributions of High (H), Mid (M), Low (L) groups of molecules

For 30S binding to R-mRNA^{+30}, three types of 30S binding time distribution were observed by plotting an accumulated distribution of total binding time of 30S for all the conditions pooled together (Figure 4.13). The cumulative binding time histograms were fitted with three Gaussian plots (with R-Sq value of 0.9914), showing there are three types of total binding populations, which were assigned as H, M, and L to represent the binding regime they cover. The population H represented the high-range of binding time of the 30S (>30% of the total observation window), M represented mid-range binding time (between 20-30% of the total observation window) and L represented low-range binding time (<10% of the observational window). These regions for H, M, and L were used as cutoffs for calculating cumulative or percentage population for each condition. The cumulative or percentage population for each regime was calculated for each condition by counting the total binning for each regime.



4.5.9 Figure 4.13 Cumulative plot of all histograms pooled together to identify distribution of bound time populations raster plot distributions

For determining the nature and number of standby (short) and initiation-active (long) binding events, we represented a random selection of 100 molecules for each condition in a raster plot[149,224]. First, the molecules were clustered into H, M, and L groups by analyzing the total binding time of the 30S each molecule categorized from the cutoffs estimated from the total binding time histogram distributions shown for each condition. To categorize the binding events into standby (red) and initiation-active (blue) binding events, we took the geometric mean of the two components of binding times obtained from the double-exponential fitting of the cumulative plot for the binding times obtained earlier. We then counted the number of red and blue events for required conditions and used them the compare any change in the nature of ‘standby’ and ‘initiation-active’ binding events under the influence of preQ$_1$ or the effect of strategic mutations. We plotted each individual molecule’s 30S binding behavior categorized into the groups of H, M, and L to represent binding events as standby or initiation-active events (shown in red or blue) within each condition.

4.5.10 RNA sequences for single-molecule and biochemical assays

RNA annotation	Sequence
R-mRNA^{+30}	GGGCAGUGAGCAACAAAAUGCUCACCUGGGUCGCAGUAACCCCAGUUAACAAAACAAGGGAGGUAAUUUUGUGCCCAAAAAAAGAAUAAAAGAUUUAGCU
I1S8	GGGCAGUGAGCAACAAAAUGCUCACCUGGGUCGCAGUAACCCCAGUUAACAAAACAAGUAGGGAGGUAAUUUUGUGCCCAAAAAAAGAAUAAAAGAUUUAGCU
I2S8	GGGCAGUGAGCAACAAAAUGCUCACCUGGGUCGCAGUAACCCCAGUUAACAAAACAAGAUAGGGAGGUAAUUUUGUGCCCAAAAAAAGAAUAAAAGAUUUAGCU
I4S8	GGGCAGUGAGCAACAAAAUGCUCACCUGGGUCGCAGUAACCCCAGUUAACAAAACAAGAUAUAGGGAGGUAAUUUUGUGCCCAAAAAAAGAAUAAAAGAUUUAGCU
I6S8	GGGCAGUGAGCAACAAAAUGCUCACCUGGGUCGCAGUAACCCCAGUUAACAAAACAAGAUAUAUAGGGAGGUAAUUUUGUGCCCAAAAAAAGAAUAAAAGAUUUAGCU
I4S7	GGGCAGUGAGCAACAAAAUGCUCACCUGGGUCGCAGUAACCCCAGUUAACAAAACAAGAUAUUGGGAGGUAAUUUUGUGCCCAAAAAAAGAAUAAAAGAUUUAGCU
I4S6	GGGCAGUGAGCAACAAAAUGCUCACCUGGGUCGCAGUAACCCCAGUUAACAAAACAAGAUAUUUGGAGGUAAUUUUGUGCCCAAAAAAAGAAUAAAAGAUUUAGCU

I4S5	GGGCAGUGAGCAACAAAAUGCUCACCUGGGUCGCAGUAACCCCAGUUAACAAAACAAGAUAUUUUGAGGUAAUUUUGUGCCCAAAAAAAGAAUAAAAGAUUUAGCU
I4S0	GGGCAGUGAGCAACAAAAUGCUCACCUGGGUCGCAGUAACCCCAGUUAACAAAACAAGAUAUAUAUAUAUAAUUUUGUGCCCAAAAAAAGAAUAAAAGAUUUAGCU
Capture Strand	5’-GCATTTTGTTGCTCACTGCCC-biotin-3’
Capture Strand with Cy3	5’-Cy3-GCATTTTGTTGCTCACTGCCC-biotin-3’
30S labeling probe	5' – Cy5- GGG AGA TCA GGA TA -Cy5 3'

4.5.11 DNA primer sequences for site-directed mutagenesis to prepare R-mRNA mutants

Primer Sample	**Primer Sequence (5’ to 3’)**
Tte_Forward primer	/5Phos/ GAGGTAATTTTGTGCCC
Tte_I1S8 Reverse Primer	/5Phos/ CCTACTTGTTTTGTTAACTGG
Tte_I2S8 Reverse Primer	/5Phos/ CCTATCTTGTTTTGTTAACTGG
Tte_I4S8 Reverse Primer	/5Phos/ CCTATATCTTGTTTTGTTAACTGG
Tte_I6S8 Reverse Primer	/5Phos/ CCTATATATCTTGTTTTGTTAACTGG
Tte_I4S7 Reverse Primer	/5Phos/ CCAATATCTTGTTTTGTTAACTGG
Tte_I4S6 Reverse Primer	/5Phos/ CAAATATCTTGTTTTGTTAACTGG
Tte_I4S5 Forward Primer	/5Phos/ CCAAAAAAAGAATAAAAGATTTAGC
Tte_I4S5 Reverse Primer	/5Phos/ GCACAAAATTACCTCAAAATATCTTG
Tte_I4S0 Reverse Primer	/5Phos/ GCACAAAATTATATATATATATATCTTGTTTTGTTAAC

Chapter 5. Conclusions and Outlooks

5.1 Overview

Bacterial resistance to modern antibiotics poses one of the greatest global public health challenges of our time[225]. Pharmaceutical and drug development industries have had serious difficulties meeting this challenge due to high development costs and diminishing treatment options[226]. Hence it is essential to develop innovative strategies to target bacterial cellular processes. Traditionally, most antibiotics work by targeting ribosomes and interfering with bacterial protein synthesis[227]. However, recent advances in targeting riboswitches through artificial ligand analogs to disrupt a selective metabolic pathway in bacteria leading to its demise have opened the door for riboswitches as a model for next-generation antibiotics[228,229]. However, one of the challenges that still exists is to identify these artificial ligand analogs. This requires not just the understanding of the structural dynamics of the aptamer but also the structure-function relationship between the riboswitch aptamer and expression platform that ultimately control gene regulatory pathways in bacteria. Understanding these processes will be essential for identifying and designing potential artificial ligand analogs for specific riboswitch targets that compete with their response to cognate ligands.

The doctoral work presented here focuses on dissecting the ligand recognition and gene regulation mechanisms of two riboswitches that control gene expression by regulating transcription and translation. One of the riboswitches was observed to determine the fate of RNAP-driven transcription elongation by altering the aptamer structure through sensing Mn^{2+} ions. The other riboswitch was found to control ribosome binding and translation initiation by altering the availability of the ribosome binding site through sensing a $preQ_1$ metabolite molecule. By combining a range of single-molecule and biochemical techniques, the present work has enabled the understanding of this enigmatic gene regulatory machinery at the single-molecule level. With a robust understanding of the biochemical mechanisms of ligand binding and gene regulation based on riboswitch action at the molecular level, this work paves a broader path for understanding the central role of non-coding RNAs in bacterial gene regulation and their potential as drug targets in and beyond antibiotic resistance.

5.2 smFRET revealed that local binding of a Mn^{2+} ion to the riboswitch can trigger a large-scale response to achieve transcription control

The Mn^{2+} sensing riboswitches regulate the expression of Mn^{2+} ions homeostasis genes that decide the concentration of Mn^{2+} ions required for the survival of bacterial. They alter a part of their 4WJ structure between terminator and anti-terminator upon sensing Mn^{2+} ions at sub millimolar concentration. While the structures of these riboswitches with and without ligand were previously known[106], their mechanism of action for regulating transcription was unknown. By implementing smFRET and SiM-KARTS in combination with techniques like x-ray crystallography and MD simulations, we established the mechanism of Mn^{2+} ion sensing by this riboswitch from a rice pathogen, *X. oryzae,* and its role in regulating transcription[82]. The crystal structures for the Mn^{2+}-free and bound riboswitch aptamers resolved at 2.96 Å showed two ligand-binding cores (L1 and L3) that allowed specific binding of Mg^{2+} and Mn^{2+} based on their coordination spheres. MD simulations revealed that a stable conformation of L1, P1.1, and L1-L3 A-minor interactions are necessary for stable docking of the riboswitch. Using smFRET, we proved that stable docking was achieved only in the presence of Mn^{2+} ions, which stabilized helix P1.1 through L1-L3 interaction. Our results showed that, while Mg^{2+} binding to L1 poised the riboswitch towards a docked conformation, stable docking happened only upon binding of the Mn^{2+} ion to L3. Mutation of Mn^{2+} sensitive base A48 in L3 further supported the notion that disrupting L3 ultimately destabilized P1.1 and prevented stable docked conformation. A further test for the sensitivity of L3 to allow binding of other divalent ions like Mg^{2+}, Ni^{2+}, Co^{2+}, Zn^{2+}, $Sr^{2+,}$ and Cd^{2+}, showed that L3 is highly sensitive to sensing Mn^{2+} and triggering conformation change to upregulate transcription. This riboswitch's ability to selectively choose Mn^{2+} over other divalent ions in the cellular environment makes it an intriguing gene regulatory system. SiM-KARTS assay further revealed that stabilization of P1.1 region necessary for upoOther recent smFRET studies with a Mn^{2+} riboswitch from *L. lactis* have identified a similar mechanism that corroborates our kinetic model[124].

Such a metal ion sensitive structural rearrangement signaling transcription control is not restricted to Mn^{2+} sensing riboswitches and has been observed for other types of riboswitches like

Mg^{2+} sensing (M-box) riboswitch[104], NiCo riboswitch[125,154], and Fe^{2+} sensing riboswitch[230,231]. A recent smFRET study on a transcription regulating NiCo riboswitch from *E. bacterium* showed the highly dynamic nature of the RNA featuring a 4WJ structure that dynamically fluctuates to sample three stable conformations (unfolded, folded, and intermediate) at different concentrations and ratios of Ni^{2+}/Co^{2+} ions[154]. Such sequential folding via an intermediate conformation over simple two-state folding can be crucial for selective cognate ligand recognition and ligand-induced folding[154]. In fact, a deeper dive into some of the traces that we collected from the Mn^{2+} sensing riboswitch at different concentrations of Mg^{2+} revealed occasional sampling of a similar intermediate conformation, although dwarfed by the two predominant conformations (of FRET values 0.2 and 0.67) (Figure 5.1). The occurrence of these intermediate states indicates that riboswitches with similar 4WJ based structures may have more than two states switching, allowing the riboswitch to fold in steps for accommodating selective binding of the divalent ions. Such multi-state transitions between FRET states is not unique to metal ion binding riboswitches. Recent smFRET studies with a SAM-I riboswitch aptamer forming a similar 4WJ have also revealed dynamic transitions through four conformational states to bind SAM ligand[47].

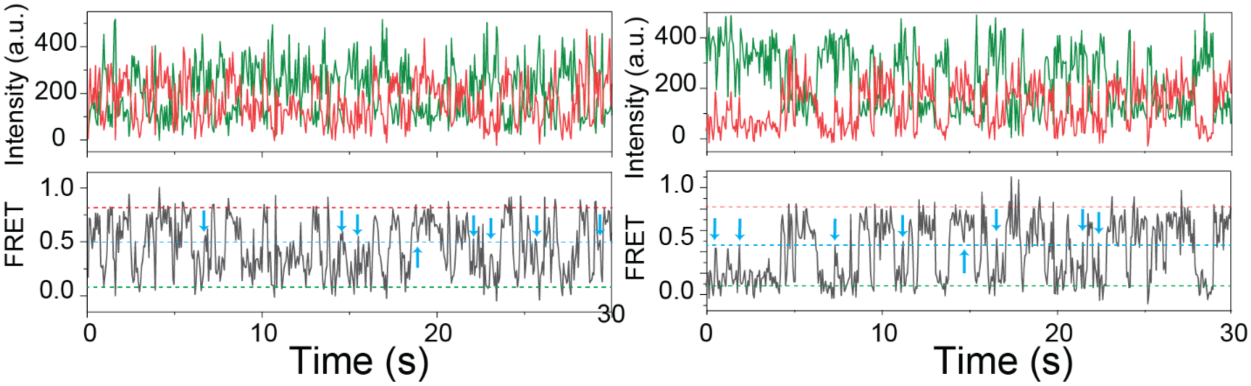


Figure 5.1 Single-molecule time traces for Mn^{2+}-sensing riboswitch from *X. oryzae* recorded at 1 mM Mg^{2+} ions showed two predominant dynamic FRET states at ~0.2 and ~0.7 and occasionally sampled intermediate FRET states (~0.45-0.5) highlighted with blue arrows.

Outlook

With our observations and results from chapter 2, we propose additional single-molecule studies to understand the stability and conformational dynamics of these intermediate states in the Mn^{2+} sensing riboswitch and their contribution to the overall signaling for gene regulation. One way to identify and characterize these short-lived intermediate states is to perform smFRET at a time resolution higher than 60 ms. Our new TIRF microscope with sCMOS camera capability (Figure 1.2) can enable single-molecule measurements at a time resolution as low as 10 ms without compromising the signal-to-noise ratio. Monitoring smFRET of the riboswitch under different divalent ion conditions at 10 ms time-resolution may allow longer observation of these short-lived intermediate states. This will also allow us to understand how the hierarchical binding of Mg^{2+} and Mn^{2+} to the ligand-binding core is associated with the stepwise folding of the riboswitch to attain stable docking.

Another single-molecule method that has been essential to understanding these intricate folding behaviors of riboswitches is single-molecule force spectroscopy using optical tweezers. Force-based single-molecule measurements have been susceptible to monitoring the hierarchical folding of riboswitch aptamers and identifying short-lived intermediate states between thermodynamically stable folded and unfolded conformations[32,232-234]. Recent developments with optical tweezers involving dual optical trapping with confocal microscopy (also commonly known as *fleezer*) have further enabled direct probing of secondary and tertiary structural changes during riboswitch folding and measuring force dependent extension and its correlation with FRET for a single molecule[235]. Simultaneous observation of FRET conformations and their influence under force in the pico-Newton (pN) force regime as well as at different concentrations of the ligand may provide unique insights into these intermediate conformational states as well as their transition kinetics.

Implementing such a combination of smFRET with force-based optical tweezer can be essential for direct measurement of intermediate conformations in Mn^{2+} sensing riboswitch. Using this technique, one can clearly detect how many conformational states are there in the riboswitch, and whether these conformational states correlate with the FRET values obtained through pTIRFM based smFRET measurements. A variation of the concentration and the ratio of divalent ions (Mg^{2+} and Mn^{2+}) would further provide insights into the variation in the occurrence and transition kinetics of the intermediate states with respect to the predominant conformational states in the free energy landscape of riboswitch folding. One specific direction in which the Mn^{2+} sensing riboswitch from *X. oryzae* can be studied would be to probe the influence of force on the stability of the A-minor interaction achieved due to divalent metal ion assisted L1-L3 stacking. Since stabilization of the L1-L3 interaction is essential for stabilization of P1.1 upon Mn^{2+} binding and resulting in a docked conformation, we hypothesize that probing the strength of the L1-L3 A-minor interaction and its influence with divalent ions under the influence of force within pN range would be a starting point to study such riboswitches.

For such experiments, the riboswitch would need certain strategic modifications to attach fluorophores and DNA handles. In this direction, one RNA strand could be designed to carry two fluorophores while the other RNA would require chimeric DNA extensions for ligations with blunt end DNA handles. The DNA handles (left and right) can be prepared from lambda-DNA through specific digestion to attain a specific length. The ends of the DNA handle can be modified for streptavidin-biotin and dioxygenase-anti-dioxygenase interactions for attachment to beads used for optical trapping. By monitoring force-dependent conformational changes alongside its correlation with FRET through fast and slow pulling of the riboswitch using pN force range, the fleezer technique can unveil hidden conformational states and their stability under the influence of divalent ions (Figure 5.2). Understanding such hidden conformational states of the Mn^{2+} sensing riboswitch under the influence of force would be crucial not only to get a broader perspective for the free-energy landscape of riboswitch folding and ligand binding[32] but also to design further assays for studying co-transcriptional folding of the riboswitch[33]. Such assays with a combination of techniques can serve to identify previously unidentified conformational states with high precision without relying exclusively on one method.

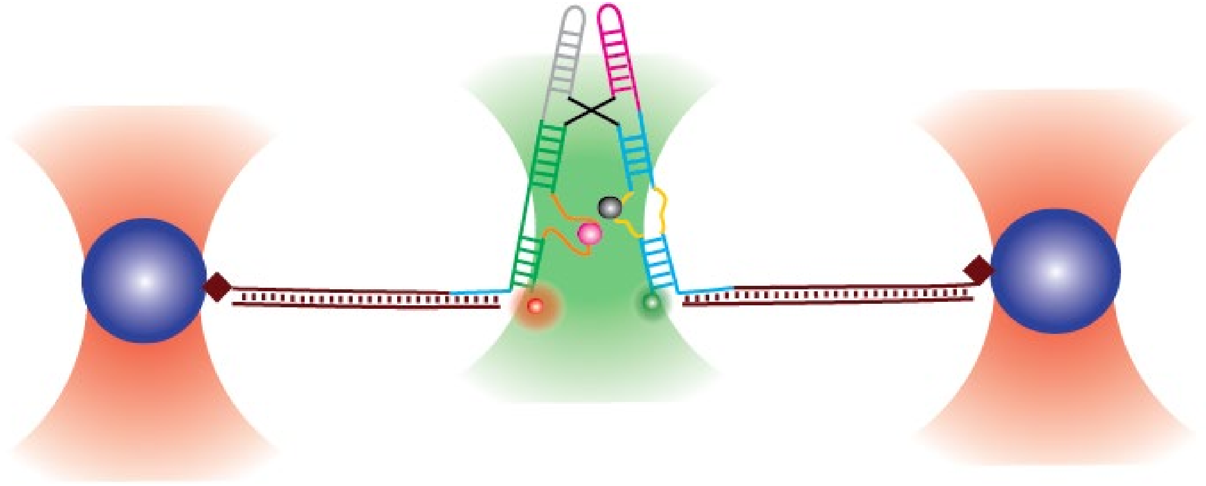


Figure 5.2 Schematic representation of the fleezer (optical trapping with smFRET) assay with fluorophore labeled Mn^{2+} sensing riboswitch attached to the beads through DNA handles on both sides.

5.3 Combination of smFRET and SiM-KARTS assay revealed the conformational dynamics of an Mn^{2+} riboswitch modulated by a paused RNA Polymerase

Chapter 2 described the mechanism of riboswitch folding in an isolated Mn^{2+} sensing riboswitch to regulate the fate of transcription. However, riboswitch folding *in vivo* usually occurs during transcription, where ligand and the activity of RNAP in terms of transcription speed and its pausing influence the folding mechanism of a riboswitch, making RNA folding a kinetically controlled process during transcription[57,236]. In addition, transcription factors like NusA have been identified to play an essential role during transcriptional pausing to govern the fate of ligand binding and riboswitch folding and hence influence the regulatory outcome for the associated gene[23,237]. To bridge this gap between transcriptional events and ligand-dependent riboswitch folding, in chapter 3, I studied the conformational dynamics of a Mn^{2+} sensing riboswitch during transcription elongation events like pausing of RNAP. Using biochemical assays, we identified a ligand-dependent class I pause of RNAP due to the riboswitch, which was used to prepare a paused elongation complex with the riboswitch associated with an RNAP through a DNA bubble. I designed single-molecule assays with the paused elongation complex (PEC) that showed that RNAP and the ligand, Mn^{2+} ions, together determine the conformational dynamics and stability of the riboswitch structure. We further identified the role of transcription factor NusA in influencing riboswitch structure and determining the fate of transcription. Our previous understanding of the Mn^{2+} sensing riboswitch and results from this chapter support a model that identified and highlighted the nuanced role of a paused RNAP elongation complex alongside Mn^{2+} ions in determining the conformational dynamics of the riboswitch and their complementary roles in regulating the outcome of transcription.

Outlook

The mechanistic details of the Mn^{2+} sensing riboswitch in the presence of a PEC provide important insights into its co-transcriptional folding and its interactions with RNAP. *In vivo* and *in vitro* studies of the Mn^{2+} responsive *yybP-ykoY* riboswitch motif have shown that it can control transcription as well as translation to regulate the expression of membrane proteins that maintain an influx of Mn^{2+} ions into the cell and regulate metal homeostasis[110,238]. Previous studies with thiamine pyrophosphate (TPP) binding riboswitches have shown similar regulation of both Rho-dependent transcription termination and repression of translation initiation by the riboswitch[137].

While recent discoveries have revealed coupling between RNA polymerase and the ribosome from *E. coli* to form an "*expressome*" complex[239], little is known about the role of the ligand-dependent riboswitch structure of the nascent mRNA on such transcription-translation coupling.

One avenue that can be explored further with the Mn^{2+} sensing riboswitch and its role in enabling transcription-translation coupling for precise gene regulation in bacteria would be to study it in the presence of an expressome complex. Such a complex would enable us to understand the role of the riboswitch in recruiting 30S ribosome subunits in the presence of an RNAP. It will also allow us to identify further the role of transcriptional pausing in determining the time window for the recruitment of the ribosomal subunit and how transcription factors such as NusA and NusG modulate 30S recruitment and thus coordinate the coupling of transcription and translation. Multicolor single-molecule assays have been implemented to understand the co-transcriptional folding of rRNA and its interaction with ribosomal proteins to study the stepwise assembly of the 30S ribosomal subunit in real-time[50,240]. From our understanding of smFRET and SiM-KARTS assays (Chapter 3 and Chapter 4) and using a combination of these single-molecule assays with the four-color pTIRF microscope (Figure 1.2), each step of the coupling into an expressome complex can be studied with intricate detail (Figure 5.3). Labeling of the riboswitch with two fluorophores (e.g., Cy2-Cy3) would determine its conformational dynamics through smFRET. Labeling of the 30S subunit with Cy5 or Alexa 647 fluorophore would evaluate the dynamics of 30S recruitment to the RBS of the mRNA. Additional fluorophore labeling of transcription factors (NusA and NusG) with near-IR fluorophores (Cy7) would determine their interaction with the expressome complex and their role in 30S recruitment and stabilization. This type of combination assays using smFRET, SiM-KARTS, and multicolor fluorophore labeling can be applied to complex transcription machinery beyond those of bacteria and be extended to eukaryotes to understand individual events over a specific time window at the single-molecule level. In combination with advancements in cryo-EM-based structure determination to understand such complex machinery as the expressome, such combinations of single-molecule assays can provide crucial mechanistic insights into the dynamics of each component in the complex that would be otherwise challenging to determine from only structural studies.

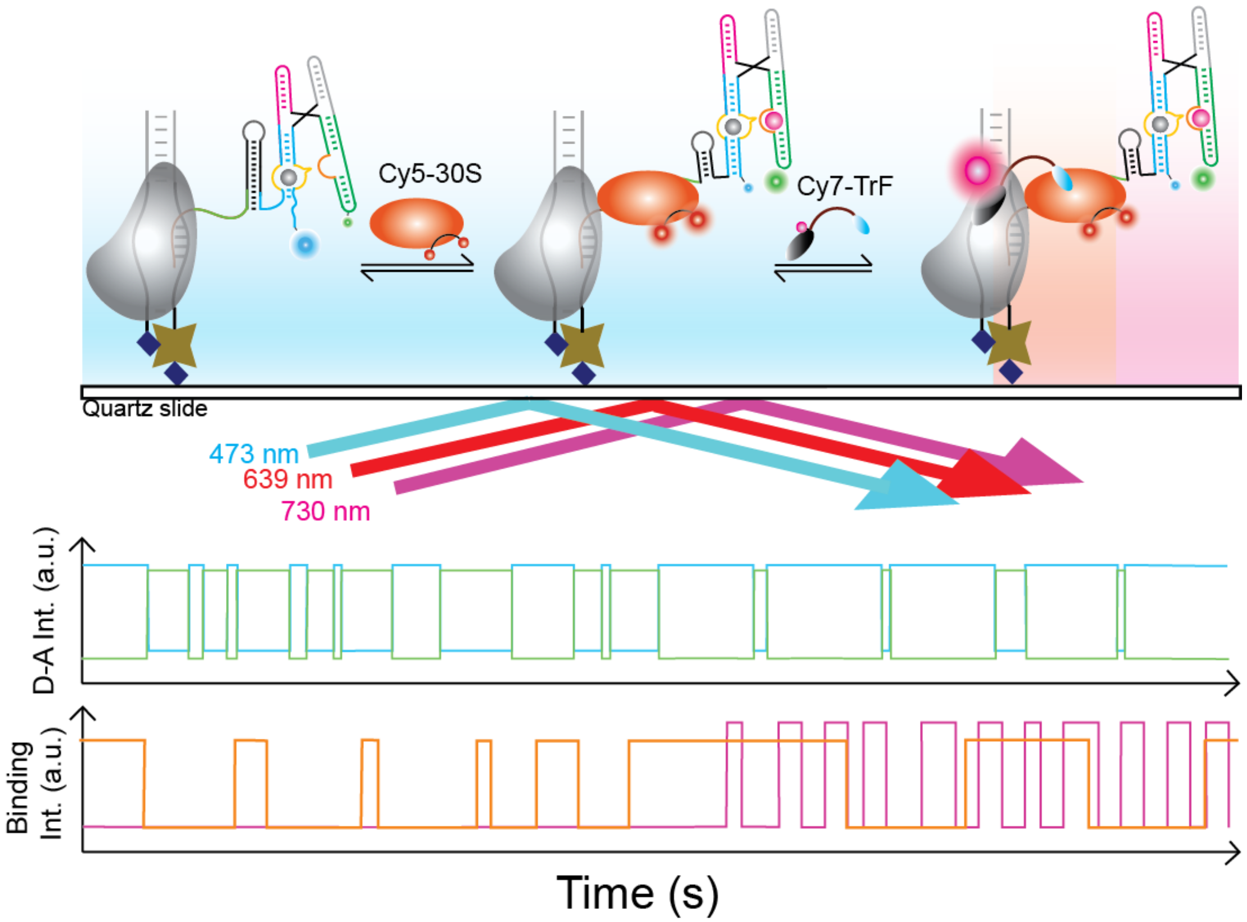


Figure 5.3 Schematic representation of a multi-color single-molecule assay to monitor transcription-translation coupling in real-time, with two representative time traces for smFRET and SiM-KARTS from a single colocalized intensities for one molecule.

The Mn^{2+} sensing riboswitch can be dual fluorophore-labeled with a Cy2-Cy3 FRET pair for monitoring conformational dynamics in response to binding Mn^{2+} ion (pink sphere, Figure 5.3). FRET changes due to conformational fluctuations are monitored through anticorrelation between the donor (blue) and acceptor (green) signal, as shown in the time trace. Once Mn^{2+} is bound, the riboswitch docks to form an anti-terminator instead of the terminator hairpin, turning transcription ON. Upon disruption of the terminator, the ribosome binding site opens and becomes accessible for Cy5-30S binding, as can be seen from the bottom trace with binding events (red). The addition of transcription factors (Cy7-TrF) during transcription to influence transcriptional pausing or facilitate transcription-translation coupling can be monitored for simultaneous colocalized binding events (pink). This method for the combination of single-molecules assays can potentially be useful to study complex RNA folding dynamics and simultaneous RNA-protein interactions through strategic fluorophore labeling of RNA and protein complexes that avoids unwanted fluorescence crosstalks. Note that the choice of color for each fluorophore and excitation sources and the representative time traces are idealized schematics.

5.4 A riboswitch in the 5'UTR is capable of determining the fate of translation in bacteria

In chapter 4, we identified a preQ_1-sensing riboswitch's role in determining the fate of binding of the 30S ribosomal subunit to initiate translation. Using a novel SiM-KARTS assay, we identified the role of ligand (preQ_1) binding, the ribosome binding site (or SD sequence), ribosomal protein S1, initiation factors with initiator tRNA, and the aptamer itself in determining the fate of translation initiation. The conformational dynamics of preQ_1 riboswitches regulating either transcription or translation have been well studied using smFRET and SiM-KARTS[22,43,65]. However, for the first time, using a combination of biochemical and SiM-KARTS assays alongside mutational analysis, I demonstrated the role of the riboswitch in the 5'-UTR of an mRNA in regulating translation control. I found that both equilibrium and non-equilibrium addition of preQ_1 profoundly affects the binding and stabilization of the 30S subunit on the mRNA. Mutational analysis of the length of the SD sequence showed that the availability of the SD sequence plays an essential role in 30S binding to mRNA. Additional mutational analysis of the distance between the aptamer and the SD sequence showed that the further the aptamer in the 5'-UTR from the SD sequence, the easier it is for the 30S subunit to bind to the SD sequence. Ribosomal protein S1 is well known for its chaperone activity in unfolding pseudoknot structures in the 5'-UTR to facilitate ribosome binding[68,69]. Titration of S1 protein showed a gradual increase in 30S binding to mRNA with an increasing concentration of S1, suggesting that S1 facilitates 30S binding and translation initiation. Overall, our data in this chapter support a model wherein the stability of 30S binding is inversely related to the free energy of unfolding of the aptamer during mRNA accommodation from the standby site to the binding cleft, enabling the ligand to reduce both 30S binding speed and kinetic stability as a two-pronged mechanism to modulate translation efficiency. Our understanding of ribosome accessibility to the 5'-UTR can be utilized to understand the nuanced role of riboswitches or any structure 5'-UTR in regulating translation and further couple RNAP and ribosome into the expressome as discussed earlier.

Outlook

As discussed in chapter 4, the modest change in riboswitch conformation upon preQ_1 binding can influence the translation initiation rate. The class I preQ_1 riboswitch is one of the smallest riboswitches and is known to be highly structurally dynamic[22,43]. While our SiM-KARTS assay was able to demonstrate the importance of only two bases to sequester or release the SD sequence for 30S binding can control riboswitch regulated translation initiation, our current assays do not capture all the mechanistic details of multiple riboswitch conformations (apo and holo) and their correlation with 30S binding. smFRET studies with translation regulating riboswitches have exquisitely demonstrated the dynamics of conformational states and their correlation with the RBS[45,130,186,241]. In one study, a combination of biochemical and smFRET assays has demonstrated the role of the cobalamin-dependent riboswitch in translation control of associated genes[242]. However, little is known about the correlation between the kinetics of riboswitch conformational change with the kinetics of ribosome binding and dissociation to the RBS.

As illustrated earlier, a combination of smFRET and SiM-KARTS assay could be an excellent approach to reveal the details of the riboswitch conformational states and their roles in determining the fate of ribosome recognition for initiating translation. For 30S binding to an mRNA with embedded preQ$_1$ riboswitch, we identified two types of ribosome recognition complexes: the standby complex and the initiation-active complex. Using a multicolor labeling strategy similar to the one discussed earlier, conformational dynamics of the preQ$_1$ riboswitch can potentially be correlated with these two types of complexes and their conformational kinetics. For example, by dual-labeling the preQ$_1$ riboswitch at the 5'-UTR for smFRET and correlating it with 30S binding through SiM-KARTS, we can pinpoint the role of each conformational state in affecting binding or dissociation of the 30S subunit and regulating initial stages of translation (Figure 5.4). In this assay, the riboswitch can be labeled with Cy2 and Cy3 using previously established labeling strategies for monitoring their conformational dynamics using smFRET[243]. The ribosome can be labeled with a Cy7 oligonucleotide as described in chapter 4. This combinatorial assay could be extremely useful in determining the duration and role of static and dynamic conformational states in recruiting and stabilizing 30S ribosomal subunits and can be potentially be extended to study other translational riboswitches and their role in regulating early stages of translation initiation.

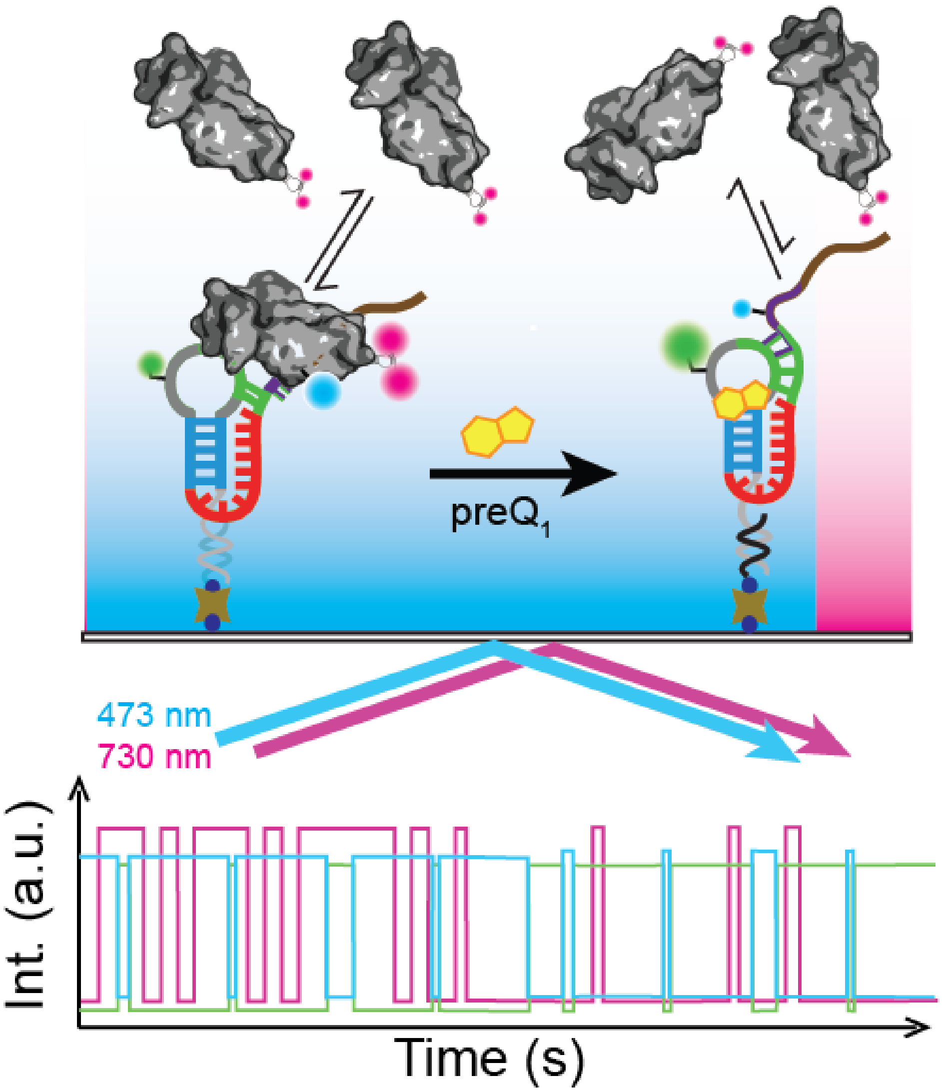


Figure 5.4 Schematic representation of a three-color single-molecule assay to monitor conformational dynamics of a preQ$_1$ riboswitch and 30S binding/dissociation at the RBS for translation initiation in real-time, with a representative time trace for smFRET and SiM-KARTS recorded for colocalized intensities from different channels for one molecule.

The preQ$_1$ sensing riboswitch can be dual fluorophore-labeled with a Cy2-Cy3 FRET pair for monitoring conformational dynamics in response to binding preQ$_1$. smFRET dynamics associated with the conformational changes can be monitored through anticorrelation between the donor (blue) and acceptor (green) signals, as shown in the representative time trace. In the absence of preQ$_1$, the riboswitch would adopt an open conformation (low FRET) and, upon the addition of preQ$_1$, a closed conformation. These conformational changes monitored from the FRET signature can be correlated with 30S binding and dissociation from the RBS. More availability of the SD sequence (more binding events in pink) can be expected during low-FRET conformational sampling, whereas less availability of the SD sequence (less binding events in pink) can be observed during the high-FRET conformational sampling. This method for combining distinct single-molecules assays can potentially be useful for studying complex RNA folding dynamics and simultaneous RNA-protein interactions through strategic fluorophore labeling of RNA and protein complexes to avoid fluorescence crosstalks. Note that the choice of color for each fluorophore and excitation source and the representative time traces are idealized schematics.

The utility of the SiM-KARTS assay to detect changes in RNA secondary structure beyond riboswitches makes it unique for a broad range of applications. With the ability to detect and probe RNA-protein interactions, one fascinating direction in which such a combined approach can be taken is the study of RNA thermometers and their roles in regulating translation initiation in bacteria. RNA thermometers are thermosensor RNA structures that undergo a conformational change in response to a shift in temperature[244]. Using such a temperature-responsive mechanism, bacteria control regulation of heat shock, bacterial pathogenesis, and virulence genes through regulating translation initiation[244-246]. While these thermal RNA switches behave like translational riboswitches, little is known about their conformational dynamics and their mechanisms of recruiting 30S subunits to initiate translation. Using a combination of smFRET and SiM-KARTS assays as described in chapter 4 and Figure 5.4, single-molecule assays can be designed to study these RNA thermometers and their mechanisms of regulation. To this end, RNA thermometers will be dual fluorophore-labeled for smFRET, and the temperature-dependent accessibility to the binding of fluorophore-labeled 30S subunit can be monitored by SiM-KARTS. The prism-TIRF microscope (Figure 1.2) allows heating of the slide holder (and the sample chamber) from room temperature to ~50(±2) °C. While designing and establishing such an assay may not be the challenging part, the thermodynamic control of ribosome binding and its dissociation from the mRNA as well as its response to the variations in temperature should be taken into consideration while designing a SiM-KARTS assay. It is anticipated that single-molecule techniques and assays like smFRET and SiM-KARTS will become increasingly valuable in understanding RNA folding and RNA-protein interactions beyond riboswitches.

Overall Summary

In this dissertation, I have studied the conformational dynamics of a Mn^{2+} sensing and a $preQ_1$ sensing riboswitch and their role in regulating transcription and translation at the single-molecule level. In chapter 1, I provided a general introduction to the riboswitches and their role in regulating transcription and translation under kinetically controlled environment inside the bacteria, and single-molecule tools to study these kinetic events. In chapter 2, I investigated the global structure and ligand-dependent conformational dynamics of an Mn^{2+} sensing riboswitch from *X. oryzae* using smFRET and corroborated them using x-ray crystallography and MD simulations. Our work unveiled a riboswitch structure with a 4WJ that samples transient docking in the absence of Mn^{2+} ions and stable docking only in the presence of Mn^{2+} ions. Our work also showed the riboswitch's specificity to bind Mn^{2+} due to its specific coordination within the ligand-binding core, which cannot stably bind other commonly available divalent ions. Our model demonstrated that the subtle binding of a small divalent ion (Mn^{2+}) to the local core of the riboswitch can result in global changes in RNA structural dynamics that upregulate transcription. With our understanding of the Mn^{2+} sensing riboswitch aptamer, in chapter 3, I further investigated the ligand-dependent conformational dynamics of the riboswitch within a paused elongation complex. Using smFRET and SiM-KARTS, I observed that RNAP, along with Mn^{2+} ions, decides the riboswitch's conformational fate. Furthermore, I also identified the role of the transcription factor, NusA, during co-transcriptional folding and conformational selection during transcriptional pausing that determines the fate of translation elongation and termination. Our overall model showed that riboswitch folding during transcription is regulated beyond just the cognate ligand and influenced by transcriptional events like pausing and binding of transcriptional factors like NusA. In chapter 4, I used our understanding of the conformational dynamics of the $preQ_1$ riboswitch from smFRET studies to determine the role of a translational $preQ_1$ riboswitch in regulating the binding of 30S ribosomal subunit to guide the fate of translation initiation. Using a SiM-KARTS assay, I studied the early stage of translation initiation near a $preQ_1$ riboswitch in the 5'UTR of a translating mRNA. I found that, during riboswitch mediated translation control, multiple factors beyond $preQ_1$ ligand determine the outcome of ribosome recognition to mRNA. The availability of the SD sequence and its distance from the aptamer determine the accessibility and stability of 30S subunit binding to initiate translation. Mutational studies showed that disruption of these factors by even a single base can alter the kinetics of ribosome recognition of mRNA. My single-molecule data further revealed that the availability of ribosomal protein S1 and its chaperone activity help unfold 5'UTR structures to promote ribosome binding. Overall, our model based on SiM-KARTS studies showed that early stages of translation initiation can be precisely tuned by a small $preQ_1$-riboswitch in the 5'UTR to affect translation outcome. I believe that our understanding of riboswitches and their role in regulating transcription and translation outcomes at the single-molecule level provides new mechanistic insights into the regulatory function of riboswitches. These insights will pave the path for studying the dynamics of the larger gene regulation machinery including the expressome, which will ultimately lead to a comprehensive understanding of bacterial gene expression and inspiration for how to target bacteria with next-generation antibiotics.

APPENDIX

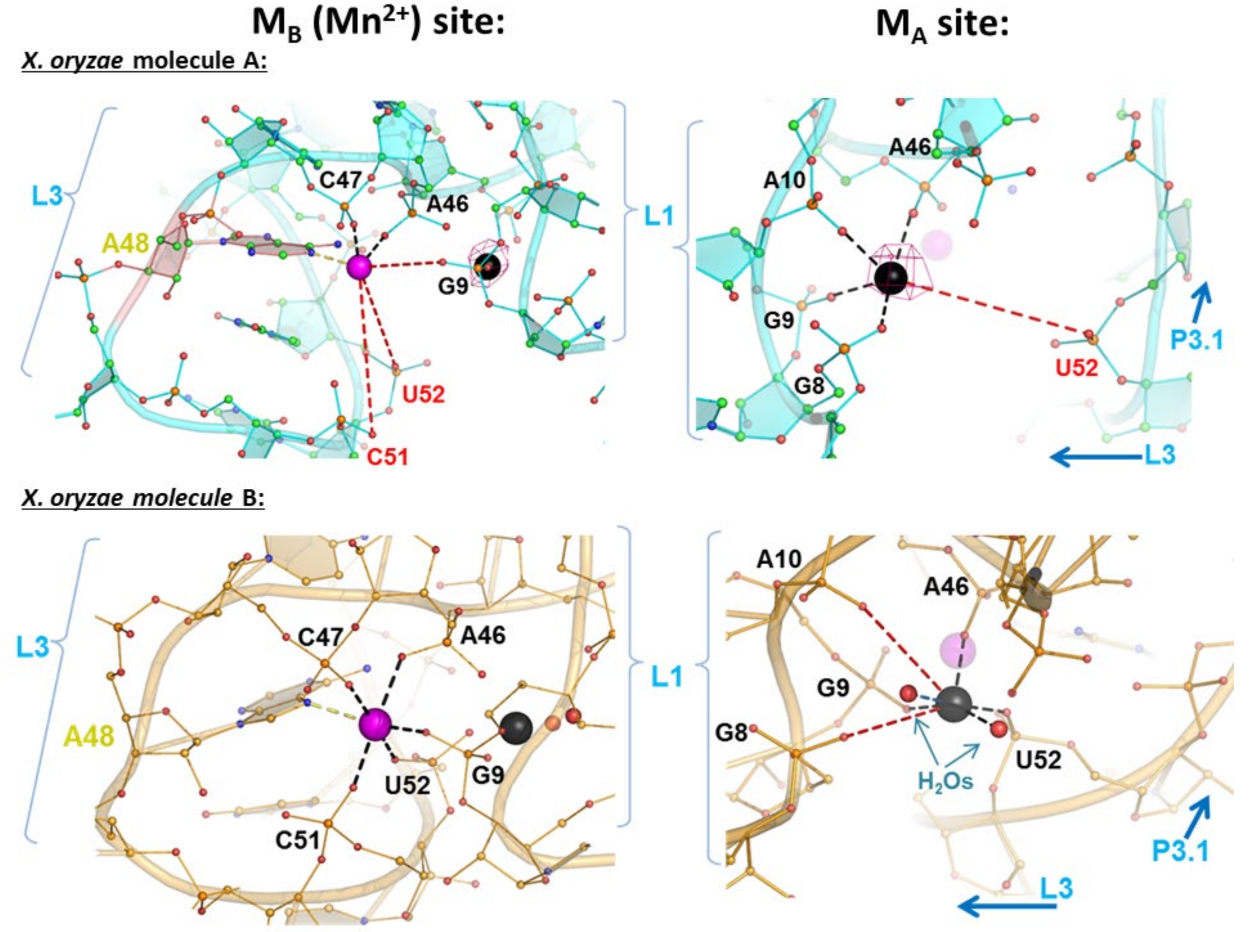


Figure 6.1 Conformational differences in the metal-binding sites in the riboswitch structures.

Close-up view of the $M_{A,Mg}$ (black) and $M_{B,Mn}$ (purple) metal-binding sites of (a) Conformer 1 and (b) Conformer 2 showing different contacts with surrounding L1 and L3 residues. Waters are shown as red spheres. The lost metal contacts are shown as red dashed lines with their distances. The anomalous difference map at 4 σ (magenta mesh) shows that Sr^{2+} is not observed in Conformer 1 and weakly occupies the $M_{A,Mg}$ site but not the $M_{B,Mn}$ site in Conformer 2.

Table 6.1 Crystallography data collection and refinement statistics. Statistics in parentheses are for the highest resolution shell.

Note that the data were mildly anisotropic, with reflections along the *l* reciprocal axis stronger. Thus, I/sigma for these reflections was well over 1.2 (indeed, over 1.5 up to 2.85 Å resolution).

Data collection:	
Wavelength (Å)	0.769
Resolution range (Å)	62.66 - 2.96 (3.07 - 2.96)
Space group	P 21 21 21
Unit cell	a/b/c: 81.32/ 85.19/ 92.49 Å α/β/γ: 90/ 90/ 90°
Total reflections	106866 (9886)
Unique reflections	13915 (1363)
Multiplicity	7.7 (7.3)
Completeness (%)	99.83 (99.63)
Mean I/sigma(I)	14.2 (1.26)
Wilson B-factor (Å^2)	81.90
R-merge	0.110 (1.51)
R-meas	0.118 (1.63)
R-pim	0.0419 (0.600)
CC1/2	1 (0.599)
CC*	1 (0.866)
Refinement:	
Reflections used in refinement	13901 (1361)
Reflections used for R-free	1136 (109)
R-work	0.205 (0.383)
R-free	0.220 (0.402)
CC(work)	0.980 (0.712)
CC(free)	0.989 (0.715)
Number of non-hydrogen atoms	4341
macromolecules	4214
ligands	120
solvent	7
RMS(bonds)	0.005
RMS(angles)	0.59
Clashscore	6.79
Average B-factor (Å^2)	69.84
macromolecules	69.17
ligands	93.43
solvent	66.69
Cruikshank DPI (Å)	0.365
Coordinate error (max. likelihood estimate) (Å)	0.40

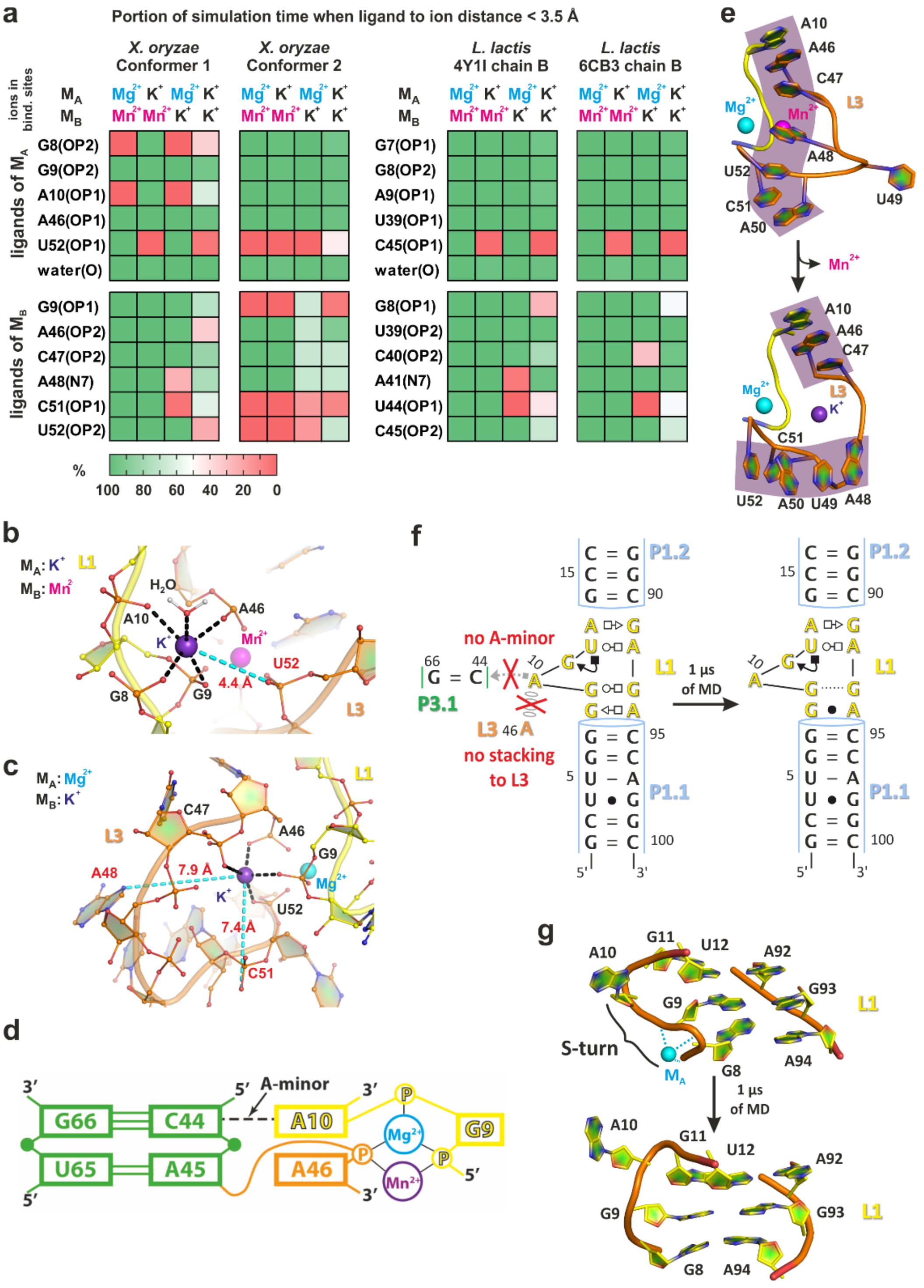


Figure 6.2 Structural dynamics of inner-shell ligands of the $M_{A,Mg}$ and $M_{B,Mn}$ sites, L3 and L1 loops as revealed by MD simulations

(**a**) The population of native inner-shell contacts in the $M_{A,Mg}$ and $M_{B,Mn}$ sites in percent as a function of the type of ion in these sites. A representative structure of the (**b**) $M_{A,Mg}$ and (**c**) $M_{B,Mn}$

site as typically observed in MD simulations with a K^+ ion occupying the site. The most labile inner-shell contacts are depicted in cyan. (**d**) 2D representation of the most stable part of the L1-L3 tertiary interaction that is likely anchored by the A-minor interaction and related A10|A46 stacking. (**e**) Partial unfolding of the stacking pattern in the L3 loop observed upon replacement of Mn^{2+} with K^+ in the $M_{B,Mn}$ site. (**f**) The loss of native G9-G93 *trans* Watson-Crick/Hoogsteen and G8-A94 *trans* Sugar-Edge/Hoogsteen base pairing, and (**g**) loss of the S-turn backbone conformation involving a phosphate notch forming part of the $M_{A,Mg}$ ion binding site in response to loss of the L1-L3 tertiary interaction with its A10…G66=C44 A-minor interaction.

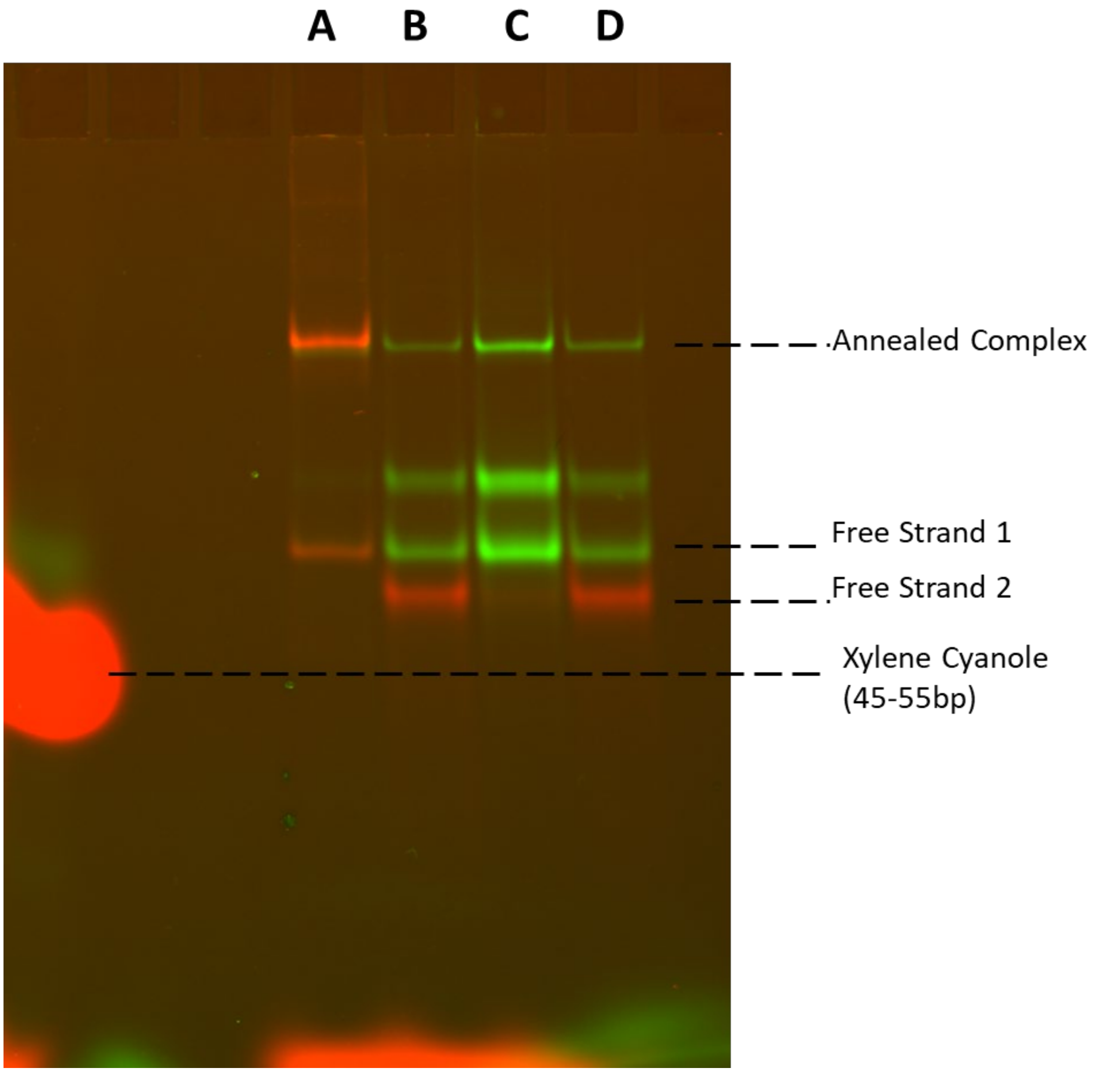


Figure 6.3 ***Xory*** **bimolecular FRET construct refolds observed from Typhoon imaging**

In PCR tube strip, samples were heated to 70 °C for 2 mins, then 2 mM $MgCl_2$ was added (or not), samples were removed and allowed to cool to RT for 10 mins, then aliquots of put on ice until loaded to gel. ~1 pmol each oligo (1-2 uL) was added to 10 uL running buffer + 10% glycerol and loaded to 15% non-denaturing PAGE gel at 4 °C. The gel wes run with 0.5X TBE + 2 mM

$MgOAc_2$ + 50 mM KCl in gel and running buffer. Lane A represents two strand annealed complex. Lane B-D represents WT strand 1 with P1.1 mutated in strand 2 (MP1.1), showing no annealing upon Pl.1 disruption.

Samples (A) 1:1 old annealed strand 1 + strand 2 in HEPES buffer annealed without Mg^{2+} in the solution (B) 1:1 new annealed strand 1 + MP1.1 strand 2 in HEPES buffer with 2 mM $MgCl_2$ (C) 1:5 new annealed strand 1 + WT MP1.1 strand 2 in HEPES buffer annealed with 2 mM $MgCl_2$ (D) 1:1 new annealed strand 1 + WT MP1.1 strand 2 in NSC buffer annealed without Mg^{2+}.

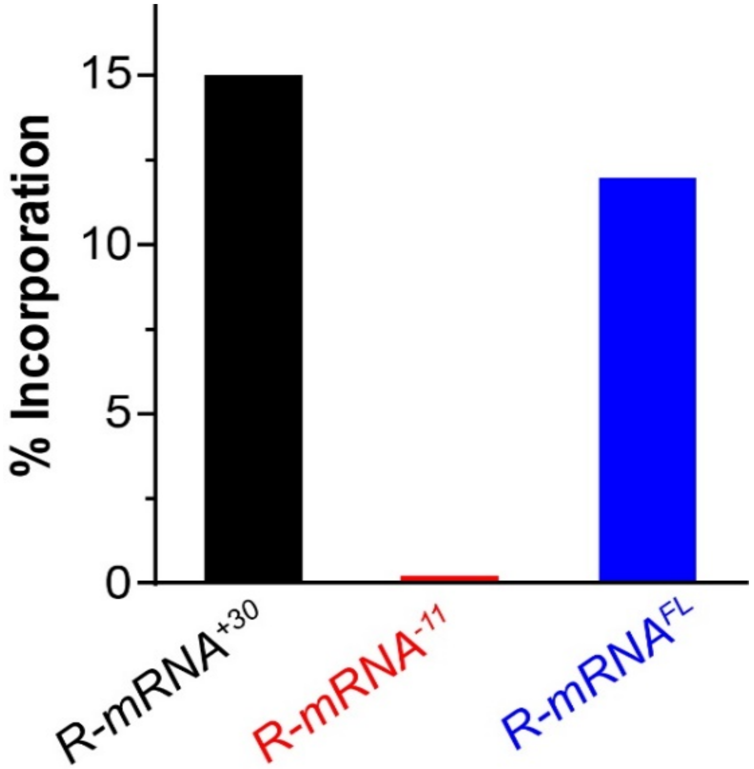


Figure 6.4 R-mRNA truncations to determine RNA required for 30S IC formation

Comparison of 30S IC formation efficiency on full length R-mRNA^{FL} and R-mRNA^{+30} truncation and R-mRNA^{-11} truncation (no SD/ ORF). As expected, ICs do not form on mRNA that lacks a SD and ORF, whereas the efficiency of initiation is similar for full-length R-mRNA and R-mRNA^{+30}.

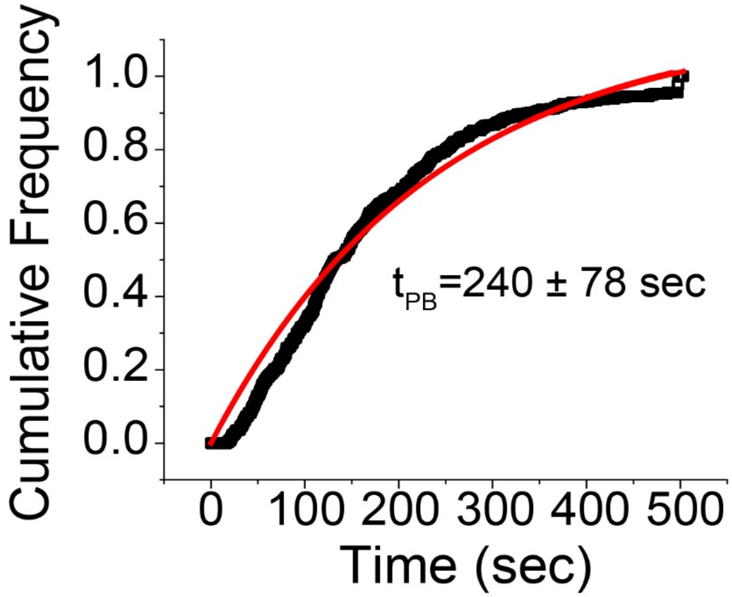


Figure 6.5 Characterization of photobleaching time for Cy5 fluorophore used for 30S labeling

The dual Cy5 labeled 30S labeling-oligo was captured on the surface by a biotinylated capture strand. Same assay conditions and laser intensity and optical parameters were maintained as the SiM-KARB assays. The average characteristic photobleaching time is shown in the graph, which is an estimation of photobleaching of both dye molecules.

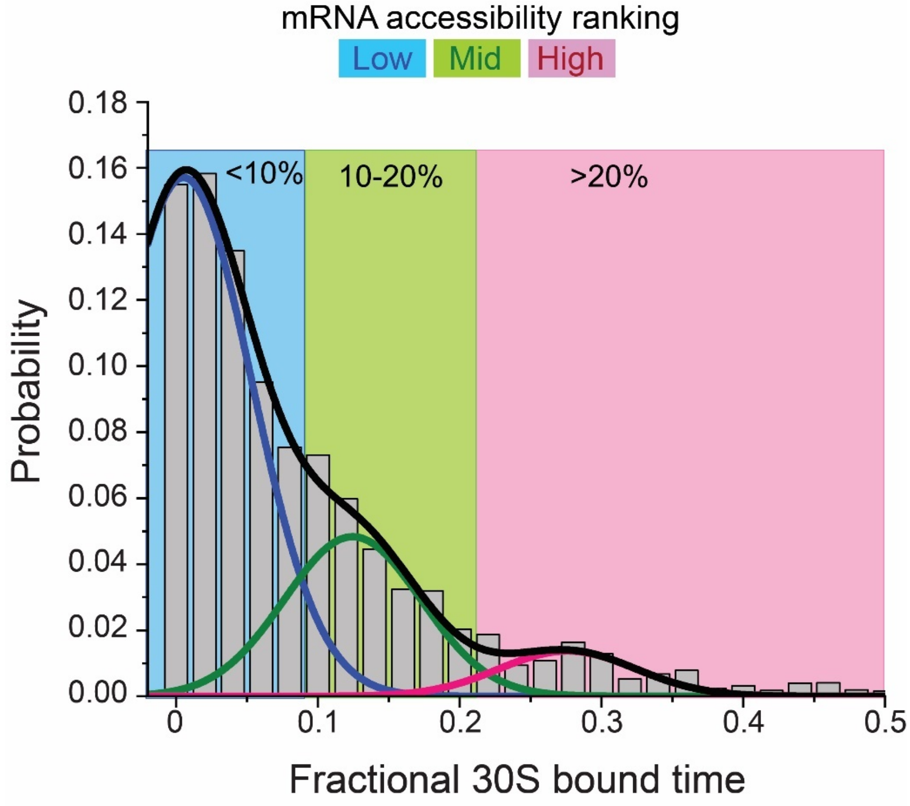


Figure 6.6 Cumulative plot of all histograms pooled together to identify distribution of bound time populations

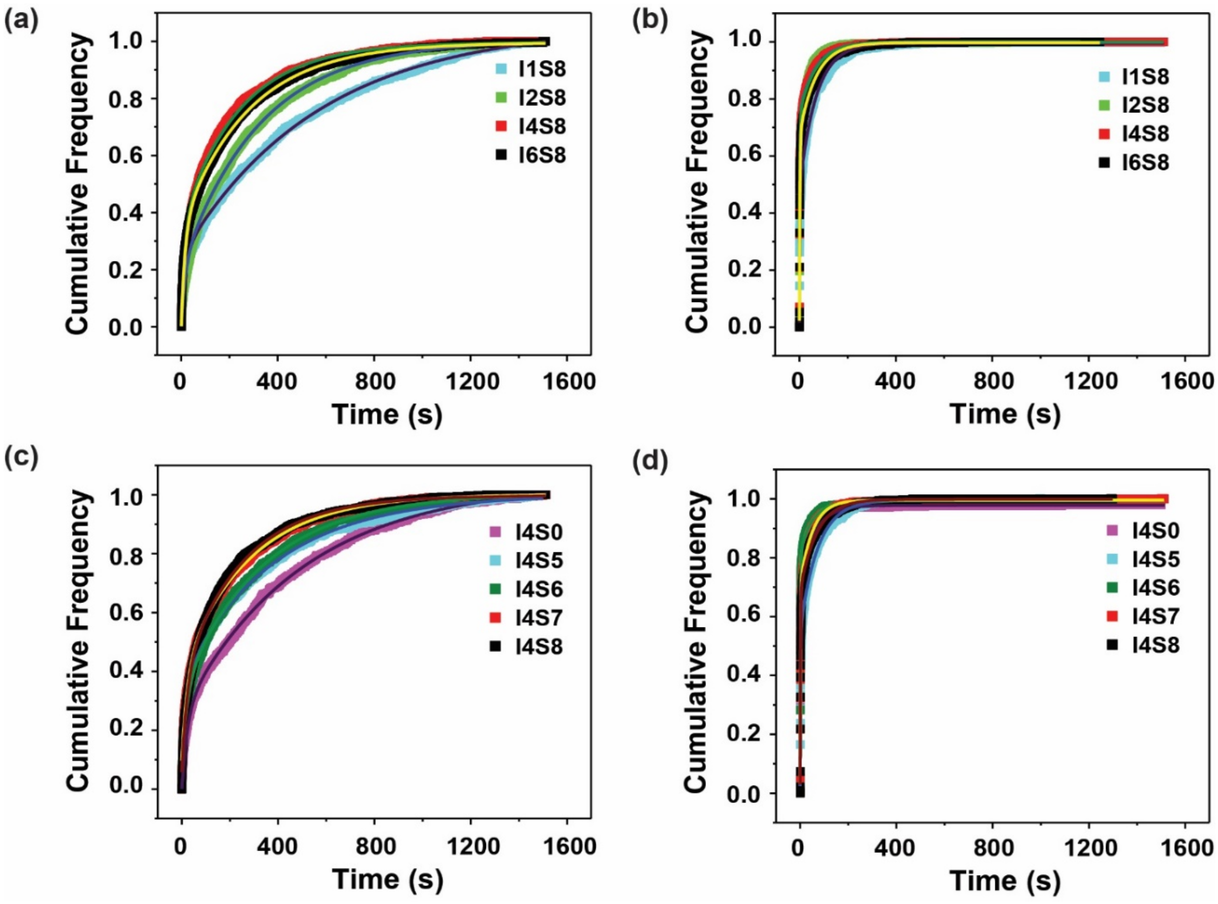


Figure 6.7 Cumulative plots for bound and unbound times for different mutants.

(a) Cumulative frequency plot for $t_{unbound}$ with different SD-aptamer distances (I1S8 to I6S8); (b) Cumulative frequency plot for t_{bound} with different SD-aptamer distances (I1S8 to I6S8). (c) Cumulative frequency plot for $t_{unbound}$ with different lengths of SD region (I4S8 to I4S0); (b) Cumulative frequency plot for t_{bound} with different lengths of SD region (I4S8 to I4S0).

Table 6.2 List of all unbound times ($t_{unbound}$) and on rates (k_{on}) for mutants with different SD-aptamer distance (I4S8 to I4S0)

	$t_{unbound, slow}$ *(s)*	±Δ $t_{unbound,slow}$ *(s)*	A_1	±ΔA_1	$k_{on,slow}$ ($\times 10^6 M^{-1}s^{-1}$)	±Δ $k_{on,slow}$ ($\times 10^6 M^{-1}s^{-1}$)	$k_{on,slow}$ *(PB corrected)* ($\times 10^6 M^{-1}s^{-1}$)	±Δ $k_{on,slow}$ *(PB corrected)* ($\times 10^6 M^{-1}s^{-1}$)
I4S8	*248.1*	*6.7*	*0.63*	*0.02*	*0.202*	*0.006*	*0.359*	*0.009*
I4S7	*284.3*	*6.3*	*0.60*	*0.02*	*0.176*	*0.004*	*0.334*	*0.007*
I4S6	*342.1*	*17.2*	*0.67*	*0.03*	*0.146*	*0.007*	*0.304*	*0.015*
I4S5	*369.4*	*30.4*	*0.63*	*0.01*	*0.135*	*0.011*	*0.297*	*0.024*
I4S0	*484.1*	*35.5*	*0.81*	*0.02*	*0.103*	*0.008*	*0.262*	*0.019*

	$t^{shared}_{unbound,fast}$ (s)	±Δ $t^{shared}_{unbound,fast}$ (s)	A_2	±ΔA_2	$k_{on,fast}$ ($\times 10^6 M^{-1} s^{-1}$)	±Δ $k_{on,fast}$ ($\times 10^6 M^{-1} s^{-1}$)	$k_{on,fast}$ (PB corrected) ($\times 10^6 M^{-1} s^{-1}$)	±Δ $k_{on,fast}$ (PB corrected) ($\times 10^6 M^{-1} s^{-1}$)
I4S8	16.24	1.23	0.37	0.02	3.078	0.233	3.237	0.245
I4S7	16.24	1.23	0.40	0.02	3.078	0.233	3.237	0.245
I4S6	16.24	1.23	0.34	0.02	3.078	0.233	3.237	0.245
I4S5	16.24	1.23	0.37	0.01	3.078	0.233	3.237	0.245
I4S0	16.24	1.23	0.22	0.01	3.078	0.233	3.237	0.245

Table 6.3 List of all bound times (t_{bound}) and off rates (k_{off}) for mutants with different SD-aptamer distance (I4S8 to I4S0)

	$t_{bound, slow}$ (s)	±Δ $t_{bound,slow}$ (s)	A_1	±ΔA_1	$k_{off,slow}$ ($\times 10^6 M^{-1} s^{-1}$)	±Δ $k_{off,slow}$ ($\times 10^6 M^{-1} s^{-1}$)	$k_{off,slow}$ (PB corrected) ($\times 10^6 M^{-1} s^{-1}$)	±Δ $k_{off,slow}$ (PB corrected) ($\times 10^6 M^{-1} s^{-1}$)
I4S8	114.9	43. 1	0.32	0.02	0.009	0.003	0.004	0.001
I4S7	53.62	4.1	0.30	0.02	0.019	0.001	0.014	0.001
I4S6	57.83	11.2	0.17	0.02	0.017	0.003	0.012	0.002
I4S5	131.7	42.6	0.45	0.03	0.008	0.002	0.003	0.001
I4S0	64.5	6.9	0.24	0.02	0.016	0.002	0.011	0.001

	$t^{shared}_{bound,fast}$ (s)	±Δ $t^{shared}_{bound,fast}$ (s)	A_2	±ΔA_2	$k_{off,fast}$ ($\times 10^6 M^{-1} s^{-1}$)	±Δ $k_{off,fast}$ ($\times 10^6 M^{-1} s^{-1}$)	$k_{off,fast}$ (PB corrected) ($\times 10^6 M^{-1} s^{-1}$)	±Δ $k_{off,fast}$ (PB corrected) ($\times 10^6 M^{-1} s^{-1}$)
I4S8	1.72	0.20	0.67	0.02	0.582	0.067	0.577	0.067
I4S7	1.72	0.20	0.70	0.02	0.582	0.067	0.577	0.067
I4S6	1.72	0.20	0.82	0.02	0.582	0.067	0.577	0.067
I4S5	1.72	0.20	0.54	0.03	0.582	0.067	0.577	0.067
I4S0	1.72	0.20	0.73	0.02	0.582	0.067	0.577	0.067

Table 6.4 List of all unbound times ($t_{unbound}$) and on rates (k_{on}) for mutants with different SD-aptamer distance (I1S8 to I6S8)

	$t_{unbound, slow}$ (s)	±Δ $t_{unbound,slow}$ (s)	A_1	±ΔA_1	$k_{on,slow}$ ($\times 10^6 M^{-1} s^{-1}$)	±Δ $k_{on,slow}$ ($\times 10^6 M^{-1} s^{-1}$)	$k_{on,slow}$ (PB corrected) ($\times 10^6 M^{-1} s^{-1}$)	±Δ $k_{on,slow}$ (PB corrected) ($\times 10^6 M^{-1} s^{-1}$)
I1S8	558.7	55.6	0.79	0.02	0.089	0.009	0.248	0.025
I2S8	327.4	11.2	0.79	0.02	0.153	0.005	0.311	0.011
I4S8	247.3	9.4	0.62	0.01	0.202	0.008	0.361	0.014

I6S8	*277.7*	*10.6*	*0.64*	*0.01*	*0.180*	*0.007*	*0.338*	*0.013*

	$t^{shared}_{unbound,fast}$ *(s)*	±Δ $t^{shared}_{unbound,fast}$ *(s)*	A_2	±ΔA_2	$k_{on,fast}$ (×$10^6 M^{-1}s^{-1}$)	±Δ $k_{on,fast}$ (×$10^6 M^{-1}s^{-1}$)	$k_{on,fast}$ (PB corrected) (×$10^6 M^{-1}s^{-1}$)	±Δ $k_{on,fast}$ (PB corrected) (×$10^6 M^{-1}s^{-1}$)
I1S8	*16.5*	*1.53*	*0.25*	*0.01*	*3.034*	*0.283*	*3.189*	*0.296*
I2S8	*16.5*	*1.53*	*0.21*	*0.01*	*3.034*	*0.283*	*3.189*	*0.296*
I4S8	*16.5*	*1.53*	*0.37*	*0.01*	*3.034*	*0.283*	*3.189*	*0.296*
I6S8	*16.5*	*1.53*	*0.35*	*0.01*	*3.034*	*0.283*	*3.189*	*0.296*

Table 6.5 List of all bound times (t_{bound}) and off rates (k_{off}) for mutants with different SD-aptamer distance (I1S8 to I6S8)

	$t_{bound,slow}$ *(s)*	±Δ $t_{bound,slow}$ *(s)*	A_1	±ΔA_1	$k_{off,slow}$ (×$10^6 M^{-1}s^{-1}$)	±Δ $k_{off,slow}$ (×$10^6 M^{-1}s^{-1}$)	$k_{off,slow}$ (PB corrected) (×$10^6 M^{-1}s^{-1}$)	±Δ $k_{off,slow}$ (PB corrected) (×$10^6 M^{-1}s^{-1}$)
I1S8	*84.8*	*20.2*	*0.48*	*0.02*	*0.012*	*0.003*	*0.007*	*0.002*
I2S8	*58.2*	*8.0*	*0.33*	*0.03*	*0.017*	*0.002*	*0.012*	*0.002*
I4S8	*99.3*	*37.5*	*0.30*	*0.01*	*0.010*	*0.004*	*0.005*	*0.002*
I6S8	*74.5*	*42.8*	*0.19*	*0.07*	*0.013*	*0.008*	*0.008*	*0.005*

	$t^{shared}_{bound,fast}$ *(s)*	±Δ $t^{shared}_{bound,fast}$ *(s)*	A_2	±ΔA_2	$k_{off,fast}$ (×$10^6 M^{-1}s^{-1}$)	±Δ $k_{off,fast}$ (×$10^6 M^{-1}s^{-1}$)	$k_{off,fast}$ (PB corrected) (×$10^6 M^{-1}s^{-1}$)	±Δ $k_{off,fast}$ (PB corrected) (×$10^6 M^{-1}s^{-1}$)
I1S8	*2.3*	*0.85*	*0.52*	*0.02*	*0.435*	*0.162*	*0.429*	*0.159*
I2S8	*2.3*	*0.85*	*0.66*	*0.03*	*0.435*	*0.162*	*0.429*	*0.159*
I4S8	*2.3*	*0.85*	*0.70*	*0.01*	*0.435*	*0.162*	*0.429*	*0.159*
I6S8	*2.3*	*0.85*	*0.81*	*0.07*	*0.435*	*0.162*	*0.429*	*0.159*

Table 6.6 Unbound times ($t_{unbound}$) and association rates (k_{on}) for the influence of preQ$_1$ on mutant I1S8

	$t_{unbound,slow}$ (s)	±Δ $t_{unbound,slow}$ (s)	A_1	±ΔA_1	$k_{on,slow}$ (×10^{6}M^{-1}s^{-1})	±Δ $k_{on,slow}$ (×10^6 M^{-1}s^{-1})	$k_{on,slow}$ (PB corrected) (×10^{6}M^{-1}s^{-1})	±Δ $k_{on,slow}$ (PB corrected) (×10^{6}M^{-1}s^{-1})
I1S8 -preQ$_1$	*330.9*	*61.0*	*0.51*	*0.03*	*0.151*	*0.028*	*0.309*	*0.057*
I1S8 +preQ$_1$	*290.9*	*46.1*	*0.52*	*0.01*	*0.172*	*0.027*	*0.330*	*0.052*

	$t^{shared}_{unbound,fas}$ (s)	±Δ $t^{shared}_{unbound,fast}$ (s)	A_2	±ΔA_2	$k_{on,fast}$ (×10^6 M^{-1}s^{-1})	±Δ $k_{on,fast}$ (×10^6 M^{-1}s^{-1})	$k_{on,fast}$ (PB corrected) (×10^{6}M^{-1}s^{-1})	±Δ $k_{on,fast}$ (PB corrected) (×10^{6}M^{-1}s^{-1})
I1S8 -preQ$_1$	*7.8*	*1.63*	*0.49*	*0.028*	*6.389*	*1.331*	*6.569*	*1.373*
I1S8 +preQ$_1$	*7.8*	*1.63*	*0.48*	*0.014*	*6.389*	*1.331*	*6.569*	*1.373*

Table 6.7 Bound times (t_{bound}) and dissociation rates (k_{OFF}) for the influence of preQ$_1$ on mutant I1S8

	$t_{bound,\ slow}$ (s)	±Δ $t_{bound,slow}$ (s)	A_1	±ΔA_1	$k_{off,slow}$ (×10^6 M^{-1}s^{-1})	±Δ $k_{off,slow}$ (×10^6 M^{-1}s^{-1})	$k_{off,slow}$ (PB corrected) (×10^{6}M^{-1}s^{-1})	±Δ $k_{off,slow}$ (PB corrected) (×10^{6}M^{-1}s^{-1})
I1S8 -preQ$_1$	*168.3*	*7.5*	*0.59*	*0.02*	*0.006*	*0.001*	*0.001*	*0.001*
I1S8 +preQ$_1$	*219.0*	*23.0*	*0.46*	*0.01*	*0.005*	*0.001*	*0.001*	*0.001*

	$t^{shared}_{bound,fa}$ (s)	±Δ $t^{shared}_{bound,fast}$ (s)	A_2	±ΔA_2	$k_{off,fast}$ (×10^{6}M^{-1}s^{-1})	±Δ $k_{off,fast}$ (×10^{6}M^{-1}s^{-1})	$k_{off,fast}$ (PB corrected) (×10^{6}M^{-1}s^{-1})	±Δ $k_{off,fast}$ (PB corrected) (×10^{6}M^{-1}s^{-1})
I1S8 -preQ$_1$	*5.8*	*0.51*	*0.39*	*0.02*	*0.172*	*0.015*	*0.167*	*0.015*
I1S8 +preQ$_1$	*5.8*	*0.51*	*0.53*	*0.01*	*0.172*	*0.015*	*0.167*	*0.015*

Table 6.8 Unbound times ($t_{unbound}$) and association rates (k_{on}) for the influence of preQ$_1$ on mutant construct I4S8

	$t_{unbound,slow}$ (s)	±Δ $t_{unbound,slow}$ (s)	A_1	±ΔA_1	$k_{on,slow}$ ($\times 10^6 M^{-1}s^{-1}$)	±Δ $k_{on,slow}$ ($\times 10^6 M^{-1}s^{-1}$)	$k_{on,slow}$ (PB corrected) ($\times 10^6 M^{-1}s^{-1}$)	±Δ $k_{on,slow}$ (PB corrected) ($\times 10^6 M^{-1}s^{-1}$)
I4S8 -preQ$_1$	385.0	31.4	0.54	0.01	0.129	0.011	0.288	0.024
I4S8 +preQ$_1$	403.0	20.7	0.53	0.02	0.124	0.006	0.282	0.015

	$t^{shared}_{unbound,fas}$ (s)	±Δ $t^{shared}_{unbound,f}$ (s)	A_2	±ΔA_2	$k_{on,fast}$ ($\times 10^6 M^{-1}s^{-1}$)	±Δ $k_{on,fast}$ ($\times 10^6 M^{-1}s^{-1}$)	$k_{on,fast}$ (PB corrected) ($\times 10^6 M^{-1}s^{-1}$)	±Δ $k_{on,fast}$ (PB corrected) ($\times 10^6 M^{-1}s^{-1}$)
I4S8 -preQ$_1$	9.6	1.6	0.45	0.018	5.192	0.869	5.367	0.894
I4S8 +preQ$_1$	9.6	1.6	0.48	0.020	5.192	0.869	5.367	0.894

Table 6.9 Bound times (t_{bound}) and dissociation rates (k_{OFF}) for the influence of preQ$_1$ on mutant construct I4S8

	$t_{bound,slow}$ (s)	±Δ $t_{bound,slow}$ (s)	A_1	±ΔA_1	$k_{off,slow}$ ($\times 10^6 M^{-1}s^{-1}$)	±Δ $k_{off,slow}$ ($\times 10^6 M^{-1}s^{-1}$)	$k_{off,slow}$ (PB corrected) ($\times 10^6 M^{-1}s^{-1}$)	±Δ $k_{off,slow}$ (PB corrected) ($\times 10^6 M^{-1}s^{-1}$)
I4S8 -preQ$_1$	162.9	21.5	0.51	0.03	0.006	0.001	0.001	0.0001
I4S8 +preQ$_1$	138.9	19.1	0.52	0.02	0.007	0.001	0.002	0.0003

	$t^{shared}_{bound,fast}$ (s)	±Δ $t^{shared}_{bound,fast}$ (s)	A_2	±ΔA_2	$k_{off,fast}$ ($\times 10^6 M^{-1}s^{-1}$)	±Δ $k_{off,fast}$ ($\times 10^6 M^{-1}s^{-1}$)	$k_{off,fast}$ (PB corrected) ($\times 10^6 M^{-1}s^{-1}$)	±Δ $k_{off,fast}$ (PB corrected) ($\times 10^6 M^{-1}s^{-1}$)
I4S8 -preQ$_1$	4.1	0.74	0.49	0.03	0.245	0.045	0.239	0.043

I4S8 +preQ$_1$	*4.1*	*0.74*	*0.47*	*0.02*	*0.245*	*0.045*	*0.239*	*0.043*

Table 6.10 Unbound times ($t_{unbound}$) and associated rates (k_{ON}) for the influence of preQ$_1$ in absence and presence of initiation factors (IFs)

	$t_{unbound, slow}$ *(s)*	±Δ $t_{unbound,slow}$ *(s)*	A_1	±ΔA_1	$k_{on,slow}$ *($\times 10^6$ $M^{-1}s^{-1}$)*	±Δ $k_{on,slow}$ *($\times 10^6 M^{-1}s^{-1}$)*	$k_{on,slow}$ *(PB corrected) ($\times 10^6 M^{-1}s^{-1}$)*	±Δ $k_{on,slow}$ *(PB corrected) ($\times 10^6 M^{-1}s^{-1}$)*
-IF; -preQ$_1$	*299.9*	*35.3*	*0.89*	*0.04*	*0.167*	*0.019*	*0.325*	*0.038*
-IF; +preQ$_1$	*607.5*	*86.1*	*0.96*	*0.08*	*0.082*	*0.012*	*0.241*	*0.034*
+IF; -preQ$_1$	*220.3*	*7.2*	*0.76*	*0.03*	*0.227*	*0.007*	*0.385*	*0.013*
+IF; +preQ$_1$	*460.8*	*72.6*	*1.04*	*0.06*	*0.108*	*0.017*	*0.267*	*0.042*

	$t^{shared}_{unbound,fast}$ *(s)*	±Δ $t^{shared}_{unbound,fast}$ *(s)*	A_2	±ΔA_2	$k_{on,fast}$ *($\times 10^6$ $M^{-1}s^{-1}$)*	±Δ $k_{on,fast}$ *($\times 10^6 M^{-1}s^{-1}$)*	$k_{on,fast}$ *(PB corrected) ($\times 10^6 M^{-1}s^{-1}$)*	±Δ $k_{on,fast}$ *(PB corrected) ($\times 10^6 M^{-1}s^{-1}$)*
-IF; -preQ$_1$	*10.3*	*1.38*	*0.15*	*0.03*	*4.851*	*0.645*	*4.845*	*0.649*
-IF; +preQ$_1$	*10.3*	*1.38*	*0.22*	*0.01*	*4.851*	*0.645*	*4.845*	*0.649*
+IF; -preQ$_1$	*10.3*	*1.38*	*0.30*	*0.02*	*4.851*	*0.645*	*4.845*	*0.649*
+IF; +preQ$_1$	*10.3*	*1.38*	*0.07*	*0.01*	*4.851*	*0.645*	*4.845*	*0.649*

Table 6.11 Bound times (t_{bound}) and dissociated rates (k_{OFF}) for the influence of preQ$_1$ in absence and presence of initiation factors (IFs)

	$t_{bound, slow}$ *(s)*	±Δ $t_{bound,slow}$ *(s)*	A_1	±ΔA_1	$k_{off,slow}$ *($\times 10^6 M^{-1}s^{-1}$)*	±Δ $k_{off,slow}$ *($\times 10^6 M^{-1}s^{-1}$)*	$k_{off,slow}$ *(PB corrected) ($\times 10^6 M^{-1}s^{-1}$)*	±Δ $k_{off,slow}$ *(PB corrected) ($\times 10^6 M^{-1}s^{-1}$)*
-IF; -preQ$_1$	*137.2*	*41.3*	*0.58*	*0.08*	*0.007*	*0.002*	*0.002*	*0.001*
-IF;	*126.2*	*58.9*	*0.41*	*0.08*	*0.008*	*0.004*	*0.006*	*0.001*

+preQ$_1$								
+IF; -preQ$_1$	*98.8*	*13.9*	*0.42*	*0.04*	*0.010*	*0.001*	*0.005*	*0.001*
+IF; +preQ$_1$	*129.4*	*39.9*	*0.50*	*0.06*	*0.008*	*0.002*	*0.003*	*0.001*

	$t^{shared}_{bound,fast}$ **(s)**	**±Δ** $t^{shared}_{bound,fast}$ **(s)**	A_2	**±Δ**A_2	$k_{off,fast}$ **(×**10^6 $M^{-1}s^{-1}$**)**	**±Δ** $k_{off,fast}$ **(×**$10^6$$M^{-1}s^{-1}$**)**	$k_{off,fast}$ **(PB corrected) (×**$10^6$$M^{-1}s^{-1}$**)**	**±Δ** $k_{off,fast}$ **(PB corrected) (×**$10^6$$M^{-1}s^{-1}$**)**
-IF; -preQ$_1$	*7.2*	*1.10*	*0.36*	*0.03*	*0.138*	*0.021*	*0.133*	*0.020*
-IF; +preQ$_1$	*7.2*	*1.10*	*0.58*	*0.08*	*0.138*	*0.021*	*0.133*	*0.020*
+IF; -preQ$_1$	*7.2*	*1.10*	*0.57*	*0.04*	*0.138*	*0.021*	*0.133*	*0.020*
+IF; +preQ$_1$	*7.2*	*1.10*	*0.46*	*0.06*	*0.138*	*0.021*	*0.133*	*0.020*

REFERENCES

1. Crick, F. Central dogma of molecular biology. *Nature* **227**, 561-3 (1970).
2. Cobb, M. 60 years ago, Francis Crick changed the logic of biology. *PLoS Biol* **15**, e2003243 (2017).
3. Lodish H, B.A., Zipursky SL, et al. . Molecular Cell Biology. 4th edition. New York: W. H. Freeman; 2000. Section 10.2, Bacterial Transcription Initiation.
4. Farabaugh, P.J. The Translational Machinery. in *Programmed Alternative Reading of the Genetic Code* 5-27 (1997).
5. Li, J. & Liu, C. Coding or Noncoding, the Converging Concepts of RNAs. *Front Genet* **10**, 496 (2019).
6. Sana, J., Faltejskova, P., Svoboda, M. & Slaby, O. Novel classes of non-coding RNAs and cancer. *J Transl Med* **10**, 103 (2012).
7. Palazzo, A.F. & Lee, E.S. Non-coding RNA: what is functional and what is junk? *Front Genet* **6**, 2 (2015).
8. Cech, T.R. & Steitz, J.A. The noncoding RNA revolution-trashing old rules to forge new ones. *Cell* **157**, 77-94 (2014).
9. Li, S., Qian, T., Wang, X., Liu, J. & Gu, X. Noncoding RNAs and Their Potential Therapeutic Applications in Tissue Engineering. *Engineering* **3**, 3-15 (2017).
10. Lagana, A., Ferro, A. & Croce, C.M. Editorial: Bioinformatics of Non-Coding RNAs with Applications to Biomedicine: Recent Advances and Open Challenges. *Front Bioeng Biotechnol* **3**, 156 (2015).
11. Mustoe, A.M., Brooks, C.L. & Al-Hashimi, H.M. Hierarchy of RNA functional dynamics. *Annu Rev Biochem* **83**, 441-66 (2014).
12. Thirumalai, D., Lee, N., Woodson, S.A. & Klimov, D. Early events in RNA folding. *Annu Rev Phys Chem* **52**, 751-62 (2001).
13. Chen, S.-J. & Dill, K.A. RNA folding energy landscapes. *Proceedings of the National Academy of Sciences* **97**, 646-651 (2000).
14. Bevilacqua, P.C., Ritchey, L.E., Su, Z. & Assmann, S.M. Genome-Wide Analysis of RNA Secondary Structure. *Annu Rev Genet* **50**, 235-266 (2016).
15. Rich, A. The Era of RNA Awakening: Structural biology of RNA in the early years. *Quarterly Reviews of Biophysics* **42**, 117-137 (2009).
16. Wan, Y. et al. Genome-wide measurement of RNA folding energies. *Mol Cell* **48**, 169-81 (2012).
17. Fay, M.M., Lyons, S.M. & Ivanov, P. RNA G-Quadruplexes in Biology: Principles and Molecular Mechanisms. *J Mol Biol* **429**, 2127-2147 (2017).

18. Blount, K.F. & Breaker, R.R. Riboswitches as antibacterial drug targets. *Nat Biotechnol* **24**, 1558-64 (2006).
19. Breaker, R.R. Riboswitches and the RNA world. *Cold Spring Harb Perspect Biol* **4**(2012).
20. Chauvier, A., Cabello-Villegas, J. & Walter, N.G. Probing RNA structure and interaction dynamics at the single molecule level. *Methods* **162-163**, 3-11 (2019).
21. Poiata, E., Meyer, M.M., Ames, T.D. & Breaker, R.R. A variant riboswitch aptamer class for S-adenosylmethionine common in marine bacteria. *RNA (New York, N.Y.)* **15**, 2046-2056 (2009).
22. Suddala, K.C., Wang, J., Hou, Q. & Walter, N.G. Mg(2+) shifts ligand-mediated folding of a riboswitch from induced-fit to conformational selection. *J Am Chem Soc* **137**, 14075-83 (2015).
23. Widom, J.R. et al. Ligand Modulates Cross-Coupling between Riboswitch Folding and Transcriptional Pausing. *Mol Cell* **72**, 541-552.e6 (2018).
24. Zhang, J. Unboxing the T-box riboswitches-A glimpse into multivalent and multimodal RNA-RNA interactions. *Wiley Interdiscip Rev RNA*, e1600 (2020).
25. Sherwood, A.V. & Henkin, T.M. Riboswitch-Mediated Gene Regulation: Novel RNA Architectures Dictate Gene Expression Responses. *Annu Rev Microbiol* **70**, 361-74 (2016).
26. McCown, P.J., Corbino, K.A., Stav, S., Sherlock, M.E. & Breaker, R.R. Riboswitch diversity and distribution. *RNA* **23**, 995-1011 (2017).
27. Weinberg, Z. et al. Detection of 224 candidate structured RNAs by comparative analysis of specific subsets of intergenic regions. *Nucleic Acids Res* **45**, 10811-10823 (2017).
28. Greenlee, E.B. et al. Challenges of ligand identification for the second wave of orphan riboswitch candidates. *RNA Biol* **15**, 377-390 (2018).
29. Machtel, P., Bakowska-Zywicka, K. & Zywicki, M. Emerging applications of riboswitches - from antibacterial targets to molecular tools. *J Appl Genet* **57**, 531-541 (2016).
30. Al-Hashimi, H.M. & Walter, N.G. RNA dynamics: it is about time. *Curr Opin Struct Biol* **18**, 321-9 (2008).
31. Duesterberg, V.K., Fischer-Hwang, I.T., Perez, C.F., Hogan, D.W. & Block, S.M. Observation of long-range tertiary interactions during ligand binding by the TPP riboswitch aptamer. *Elife* **4**, e12362 (2015).
32. Savinov, A., Perez, C.F. & Block, S.M. Single-molecule studies of riboswitch folding. *Biochim Biophys Acta* **1839**, 1030-1045 (2014).
33. Frieda, K.L. & Block, S.M. Direct observation of cotranscriptional folding in an adenine riboswitch. *Science* **338**, 397-400 (2012).
34. Kang, J.Y. et al. Structural Basis for Transcript Elongation Control by NusG Family Universal Regulators. *Cell* **173**, 1650-1662.e14 (2018).
35. Sedlyarova, N. et al. Natural RNA Polymerase Aptamers Regulate Transcription in E. coli. *Mol Cell* **67**, 30-43.e6 (2017).
36. Breaker, R.R. Riboswitches and Translation Control. *Cold Spring Harb Perspect Biol* **10**(2018).

37. Ray, S., Chauvier, A. & Walter, N.G. Kinetics coming into focus: single-molecule microscopy of riboswitch dynamics. *RNA Biol* **16**, 1077-1085 (2019).

38. Watters, K.E., Strobel, E.J., Yu, A.M., Lis, J.T. & Lucks, J.B. Cotranscriptional folding of a riboswitch at nucleotide resolution. *Nat Struct Mol Biol* **23**, 1124-1131 (2016).

39. Wickiser, J.K., Cheah, M.T., Breaker, R.R. & Crothers, D.M. The Kinetics of Ligand Binding by an Adenine-Sensing Riboswitch. *Biochemistry* **44**, 13404-13414 (2005).

40. Gilbert, S.D., Stoddard, C.D., Wise, S.J. & Batey, R.T. Thermodynamic and kinetic characterization of ligand binding to the purine riboswitch aptamer domain. *J Mol Biol* **359**, 754-68 (2006).

41. Roberts, J.W., Shankar, S. & Filter, J.J. RNA Polymerase Elongation Factors. *Annual Review of Microbiology* **62**, 211-233 (2008).

42. Rode, A.B., Endoh, T. & Sugimoto, N. Crowding Shifts the FMN Recognition Mechanism of Riboswitch Aptamer from Conformational Selection to Induced Fit. *Angewandte Chemie International Edition* **57**, 6868-6872 (2018).

43. Suddala, K.C. et al. Single transcriptional and translational preQ1 riboswitches adopt similar pre-folded ensembles that follow distinct folding pathways into the same ligand-bound structure. *Nucleic Acids Res* **41**, 10462-75 (2013).

44. Suddala, K.C. & Walter, N.G. Riboswitch structure and dynamics by smFRET microscopy. *Methods Enzymol* **549**, 343-73 (2014).

45. Holmstrom, E.D., Polaski, J.T., Batey, R.T. & Nesbitt, D.J. Single-molecule conformational dynamics of a biologically functional hydroxocobalamin riboswitch. *J Am Chem Soc* **136**, 16832-43 (2014).

46. Elenko, M.P., Szostak, J.W. & van Oijen, A.M. Single-molecule imaging of an in vitro-evolved RNA aptamer reveals homogeneous ligand binding kinetics. *J Am Chem Soc* **131**, 9866-7 (2009).

47. Manz, C. et al. Single-molecule FRET reveals the energy landscape of the full-length SAM-I riboswitch. *Nat Chem Biol* **13**, 1172-1178 (2017).

48. McCluskey, K. et al. Unprecedented tunability of riboswitch structure and regulatory function by sub-millimolar variations in physiological Mg2+. *Nucleic Acids Research* **47**, 6478-6487 (2019).

49. Sung, H.-L. & Nesbitt, D.J. Single-Molecule FRET Kinetics of the Mn2+ Riboswitch: Evidence for Allosteric Mg2+ Control of "Induced-Fit" vs "Conformational Selection" Folding Pathways. *The Journal of Physical Chemistry B* **123**, 2005-2015 (2019).

50. Duss, O., Stepanyuk, G.A., Puglisi, J.D. & Williamson, J.R. Transient Protein-RNA Interactions Guide Nascent Ribosomal RNA Folding. *Cell* **179**, 1357-1369.e16 (2019).

51. Rodgers, M.L. & Woodson, S.A. Transcription Increases the Cooperativity of Ribonucleoprotein Assembly. *Cell* **179**, 1370-1381 e12 (2019).

52. Saba, J. et al. The elemental mechanism of transcriptional pausing. *Elife* **8**(2019).

53. Yakhnin, A.V., Murakami, K.S. & Babitzke, P. NusG Is a Sequence-specific RNA Polymerase Pause Factor That Binds to the Non-template DNA within the Paused Transcription Bubble. *J Biol Chem* **291**, 5299-308 (2016).
54. Kang, J.Y., Mishanina, T.V., Landick, R. & Darst, S.A. Mechanisms of Transcriptional Pausing in Bacteria. *J Mol Biol* **431**, 4007-4029 (2019).
55. Guo, X. et al. Structural Basis for NusA Stabilized Transcriptional Pausing. *Mol Cell* **69**, 816-827.e4 (2018).
56. Walter, N.G. Biological Pathway Specificity in the Cell-Does Molecular Diversity Matter? *Bioessays* **41**, e1800244 (2019).
57. Uhm, H., Kang, W., Ha, K.S., Kang, C. & Hohng, S. Single-molecule FRET studies on the cotranscriptional folding of a thiamine pyrophosphate riboswitch. *Proc Natl Acad Sci U S A* **115**, 331-336 (2018).
58. Studer, S.M. & Joseph, S. Unfolding of mRNA secondary structure by the bacterial translation initiation complex. *Mol Cell* **22**, 105-15 (2006).
59. Shine, J. & Dalgarno, L. The 3'-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc Natl Acad Sci U S A* **71**, 1342-6 (1974).
60. Steitz, J.A. & Jakes, K. How ribosomes select initiator regions in mRNA: base pair formation between the 3' terminus of 16S rRNA and the mRNA during initiation of protein synthesis in Escherichia coli. *Proc Natl Acad Sci U S A* **72**, 4734-8 (1975).
61. Rackham, O. & Chin, J.W. A network of orthogonal ribosome x mRNA pairs. *Nat Chem Biol* **1**, 159-66 (2005).
62. Saito, K., Green, R. & Buskirk, A.R. Translational initiation in E. coli occurs at the correct sites genome-wide in the absence of mRNA-rRNA base-pairing. *Elife* **9**(2020).
63. Jacob, W.F., Santer, M. & Dahlberg, A.E. A single base change in the Shine-Dalgarno region of 16S rRNA of Escherichia coli affects translation of many proteins. *Proc Natl Acad Sci U S A* **84**, 4757-61 (1987).
64. Salis, H.M., Mirsky, E.A. & Voigt, C.A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotechnol* **27**, 946-50 (2009).
65. Rinaldi, A.J., Lund, P.E., Blanco, M.R. & Walter, N.G. The Shine-Dalgarno sequence of riboswitch-regulated single mRNAs shows ligand-dependent accessibility bursts. *Nature Communications* **7**, 8976 (2016).
66. Kolb, A., Hermoso, J.M., Thomas, J.O. & Szer, W. Nucleic acid helix-unwinding properties of ribosomal protein S1 and the role of S1 in mRNA binding to ribosomes. *Proc Natl Acad Sci U S A* **74**, 2379-83 (1977).
67. Duval, M. et al. Escherichia coli ribosomal protein S1 unfolds structured mRNAs onto the ribosome for active translation initiation. *PLoS Biol* **11**, e1001731 (2013).
68. Qureshi, N.S., Bains, J.K., Sreeramulu, S., Schwalbe, H. & Furtig, B. Conformational switch in the ribosomal protein S1 guides unfolding of structured RNAs for translation initiation. *Nucleic Acids Res* **46**, 10917-10929 (2018).

69. Lund, P.E., Chatterjee, S., Daher, M. & Walter, N.G. Protein unties the pseudoknot: S1-mediated unfolding of RNA higher order structure. *Nucleic Acids Res* **48**, 2107-2125 (2020).
70. Roy, R., Hohng, S. & Ha, T. A practical guide to single-molecule FRET. *Nat Methods* **5**, 507-16 (2008).
71. Hellenkamp, B. et al. Precision and accuracy of single-molecule FRET measurements-a multi-laboratory benchmark study. *Nat Methods* **15**, 669-676 (2018).
72. Ploetz, E. et al. Forster resonance energy transfer and protein-induced fluorescence enhancement as synergetic multi-scale molecular rulers. *Sci Rep* **6**, 33257 (2016).
73. Coats, J.E., Lin, Y., Rueter, E., Maher, L.J., 3rd & Rasnik, I. Single-molecule FRET analysis of DNA binding and bending by yeast HMGB protein Nhp6A. *Nucleic Acids Res* **41**, 1372-81 (2013).
74. Wozniak, A.K., Schroder, G.F., Grubmuller, H., Seidel, C.A. & Oesterhelt, F. Single-molecule FRET measures bends and kinks in DNA. *Proc Natl Acad Sci U S A* **105**, 18337-42 (2008).
75. LeBlanc, S.J., Kulkarni, P. & Weninger, K.R. Single Molecule FRET: A Powerful Tool to Study Intrinsically Disordered Proteins. *Biomolecules* **8**(2018).
76. Mazal, H. & Haran, G. Single-molecule FRET methods to study the dynamics of proteins at work. *Curr Opin Biomed Eng* **12**, 8-17 (2019).
77. Lerner, E. et al. Toward dynamic structural biology: Two decades of single-molecule Forster resonance energy transfer. *Science* **359**(2018).
78. Mazumder, A., Lin, M., Kapanidis, A.N. & Ebright, R.H. Closing and opening of the RNA polymerase trigger loop. *Proc Natl Acad Sci U S A* **117**, 15642-15649 (2020).
79. Juette, M.F. et al. The bright future of single-molecule fluorescence imaging. *Curr Opin Chem Biol* **20**, 103-11 (2014).
80. van der Feltz, C. & Hoskins, A.A. Methodologies for studying the spliceosome's RNA dynamics with single-molecule FRET. *Methods* **125**, 45-54 (2017).
81. Schnitzbauer, J., Strauss, M.T., Schlichthaerle, T., Schueder, F. & Jungmann, R. Super-resolution microscopy with DNA-PAINT. *Nat Protoc* **12**, 1198-1228 (2017).
82. Suddala, K.C. et al. Local-to-global signal transduction at the core of a Mn(2+) sensing riboswitch. *Nat Commun* **10**, 4304 (2019).
83. Juette, M.F. et al. Single-molecule imaging of non-equilibrium molecular ensembles on the millisecond timescale. *Nat Methods* **13**, 341-4 (2016).
84. McKinney, S.A., Joo, C. & Ha, T. Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophys J* **91**, 1941-51 (2006).
85. Bronson, J.E., Fei, J., Hofman, J.M., Gonzalez, R.L., Jr. & Wiggins, C.H. Learning rates and states from biophysical time series: a Bayesian approach to model selection and single-molecule FRET data. *Biophys J* **97**, 3196-205 (2009).

86. van de Meent, J.W., Bronson, J.E., Wiggins, C.H. & Gonzalez, R.L., Jr. Empirical Bayes methods enable advanced population-level analyses of single-molecule FRET experiments. *Biophys J* **106**, 1327-37 (2014).
87. Pereira, M.J. et al. Single VS ribozyme molecules reveal dynamic and hierarchical folding toward catalysis. *J Mol Biol* **382**, 496-509 (2008).
88. Cardo, L., Karunatilaka, K.S., Rueda, D. & Sigel, R.K. Single molecule FRET characterization of large ribozyme folding. *Methods Mol Biol* **848**, 227-51 (2012).
89. Tippana, R., Chen, M.C., Demeshkina, N.A., Ferre-D'Amare, A.R. & Myong, S. RNA G-quadruplex is resolved by repetitive and ATP-dependent mechanism of DHX36. *Nat Commun* **10**, 1855 (2019).
90. Lai, W.C. et al. mRNAs and lncRNAs intrinsically form secondary structures with short end-to-end distances. *Nat Commun* **9**, 4328 (2018).
91. Lee, J., Crickard, J.B., Reese, J.C. & Lee, T.H. Single-molecule FRET method to investigate the dynamics of transcription elongation through the nucleosome by RNA polymerase II. *Methods* **159-160**, 51-58 (2019).
92. Ruehle, M.D. et al. A dynamic RNA loop in an IRES affects multiple steps of elongation factor-mediated translation initiation. *Elife* **4**(2015).
93. Caban, K., Pavlov, M., Ehrenberg, M. & Gonzalez, R.L., Jr. A conformational switch in initiation factor 2 controls the fidelity of translation initiation in bacteria. *Nat Commun* **8**, 1475 (2017).
94. Semlow, D.R., Blanco, M.R., Walter, N.G. & Staley, J.P. Spliceosomal DEAH-Box ATPases Remodel Pre-mRNA to Activate Alternative Splice Sites. *Cell* **164**, 985-98 (2016).
95. Blanco, M.R. et al. Single Molecule Cluster Analysis dissects splicing pathway conformational dynamics. *Nat Methods* **12**, 1077-84 (2015).
96. Mandal, M. & Breaker, R.R. Gene regulation by riboswitches. *Nat Rev Mol Cell Biol* **5**, 451-63 (2004).
97. Serganov, A. & Nudler, E. A decade of riboswitches. *Cell* **152**, 17-24 (2013).
98. Grundy, F.J. & Henkin, T.M. The S box regulon: a new global transcription termination control system for methionine and cysteine biosynthesis genes in Gram-positive bacteria. *Molecular Microbiology* **30**, 737-749 (1998).
99. Winkler, W., Nahvi, A. & Breaker, R.R. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **419**, 952-6 (2002).
100. Smith, K.D. et al. Structural basis of ligand binding by a c-di-GMP riboswitch. *Nat Struct Mol Biol* **16**, 1218-1223 (2009).
101. Nelson, J.W. et al. Riboswitches in eubacteria sense the second messenger c-di-AMP. *Nat Chem Biol* **9**, 834-839 (2013).
102. Sudarsan, N. et al. Riboswitches in Eubacteria Sense the Second Messenger Cyclic Di-GMP. *Science* **321**, 411-413 (2008).

103. Green, N.J., Grundy, F.J. & Henkin, T.M. The T box mechanism: tRNA as a regulatory molecule. *FEBS Lett* **584**, 318-24 (2010).

104. Dann, C.E., 3rd et al. Structure and mechanism of a metal-sensing regulatory RNA. *Cell* **130**, 878-92 (2007).

105. Wedekind, J.E., Dutta, D., Belashov, I.A. & Jenkins, J.L. Metalloriboswitches: RNA-based inorganic ion sensors that regulate genes. *J Biol Chem* **292**, 9441-9450 (2017).

106. Price, I.R., Gaballa, A., Ding, F., Helmann, J.D. & Ke, A. Mn(2+)-sensing mechanisms of yybP-ykoY orphan riboswitches. *Mol Cell* **57**, 1110-23 (2015).

107. Jones, C.P. & Ferre-D'Amare, A.R. Long-Range Interactions in Riboswitch Control of Gene Expression. *Annu Rev Biophys* **46**, 455-481 (2017).

108. Barrick, J.C.K., Winkler WC, Nahvi A, Mandal M, Collins J, Lee M, Roth A, Sudarsan N, Jona I, Wickiser JK, Breaker RR New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. *Proc Natl Acad Sci USA* **101**, 6421–6426 (2004).

109. Argaman, L. et al. Novel small RNA-encoding genes in the intergenic regions of Escherichia coli. *Curr Biol* **11**, 941-50 (2001).

110. Dambach, M. et al. The ubiquitous yybP-ykoY riboswitch is a manganese-responsive regulatory element. *Mol Cell* **57**, 1099-109 (2015).

111. Bachas, S.T. & Ferre-D'Amare, A.R. Convergent Use of Heptacoordination for Cation Selectivity by RNA and Protein Metalloregulators. *Cell Chem Biol* **25**, 962-973 e5 (2018).

112. Liang, H. et al. Identification and functional characterization of small non-coding RNAs in Xanthomonas oryzae pathovar oryzae. *BMC Genomics* **12**, 87 (2011).

113. Li, C.T., J; Mao, D; He, C. A novel manganese efflux system, YebN, is required for virulence by Xanthomonas oryzae pv. oryzae. *PLoS One* **6**, e21983 (2011).

114. Blanco, M. & Walter, N.G. Analysis of complex single-molecule FRET time trajectories. *Methods Enzymol* **472**, 153-78 (2010).

115. Zhuang, X. et al. Correlating structural dynamics and function in single ribozyme molecules. *Science* **296**, 1473-6 (2002).

116. Rueda, D. et al. Single-molecule enzymology of RNA: essential functional groups impact catalysis from a distance. *Proc Natl Acad Sci U S A* **101**, 10066-71 (2004).

117. Okumus, B., Wilson, T.J., Lilley, D.M. & Ha, T. Vesicle encapsulation studies reveal that single molecule ribozyme heterogeneities are intrinsic. *Biophys J* **87**, 2798-806 (2004).

118. Ditzler, M.A., Rueda, D., Mo, J., Hakansson, K. & Walter, N.G. A rugged free energy landscape separates multiple functional RNA folds throughout denaturation. *Nucleic Acids Res* **36**, 7088-99 (2008).

119. Fiore, J.L., Kraemer, B., Koberling, F., Edmann, R. & Nesbitt, D.J. Enthalpy-driven RNA folding: single-molecule thermodynamics of tetraloop-receptor tertiary interaction. *Biochemistry* **48**, 2550-8 (2009).

120. Greenfeld, M., Solomatin, S.V. & Herschlag, D. Removal of covalent heterogeneity reveals simple folding behavior for P4-P6 RNA. *J Biol Chem* **286**, 19872-9 (2011).

121. Solomatin, S.V., Greenfeld, M. & Herschlag, D. Implications of molecular heterogeneity for the cooperativity of biological macromolecules. *Nat Struct Mol Biol* **18**, 732-4 (2011).

122. Solomatin, S.V., Greenfeld, M., Chu, S. & Herschlag, D. Multiple native states reveal persistent ruggedness of an RNA folding landscape. *Nature* **463**, 681-4 (2010).

123. Marek, M.S., Johnson-Buck, A. & Walter, N.G. The shape-shifting quasispecies of RNA: one sequence, many functional folds. *Phys Chem Chem Phys* **13**, 11524-37 (2011).

124. Sung, H.L. & Nesbitt, D.J. Single-Molecule FRET Kinetics of the Mn(2+) Riboswitch: Evidence for Allosteric Mg(2+) Control of "Induced-Fit" vs "Conformational Selection" Folding Pathways. *J Phys Chem B* **123**, 2005-2015 (2019).

125. Furukawa, K. et al. Bacterial riboswitches cooperatively bind Ni(2+) or Co(2+) ions and control expression of heavy metal transporters. *Mol Cell* **57**, 1088-98 (2015).

126. Wakeman, C.A., Ramesh, A. & Winkler, W.C. Multiple metal-binding cores are required for metalloregulation by M-box riboswitch RNAs. *J Mol Biol* **392**, 723-35 (2009).

127. Pleiss, J.A.D., M. L.; Uhlenbeck, O. C. T7 RNA polymerase produces 5' end heterogeneity during in vitro transcription from certain templates. *RNA* **4**, 1313-7 (1988).

128. Michelotti, N., de Silva, C., Johnson-Buck, A.E., Manzo, A.J. & Walter, N.G. A bird's eye view tracking slow nanometer-scale movements of single molecular nano-assemblies. *Methods Enzymol* **475**, 121-48 (2010).

129. Qin, F. & Li, L. Model-based fitting of single-channel dwell-time distributions. *Biophys J* **87**, 1657-71 (2004).

130. Liberman, J.A. et al. Structural analysis of a class III preQ1 riboswitch reveals an aptamer distant from a ribosome-binding site regulated by fast dynamics. *Proceedings of the National Academy of Sciences* **112**, E3485-E3494 (2015).

131. Warhaut, S. et al. Ligand-modulated folding of the full-length adenine riboswitch probed by NMR and single-molecule FRET spectroscopy. *Nucleic Acids Res* **45**, 5512-5522 (2017).

132. Mayer, A., Landry, H.M. & Churchman, L.S. Pause & go: from the discovery of RNA polymerase pausing to its functional implications. *Curr Opin Cell Biol* **46**, 72-80 (2017).

133. Webster, M.W. et al. Structural basis of transcription-translation coupling and collision in bacteria. *Science* **369**, 1355-1359 (2020).

134. Weixlbaumer, A., Leon, K., Landick, R. & Darst, S.A. Structural basis of transcriptional pausing in bacteria. *Cell* **152**, 431-41 (2013).

135. Steinert, H. et al. Pausing guides RNA folding to populate transiently stable RNA structures for riboswitch-based transcription regulation. *Elife* **6**(2017).

136. Zhang, J. & Landick, R. A Two-Way Street: Regulatory Interplay between RNA Polymerase and Nascent RNA Structure. *Trends Biochem Sci* **41**, 293-310 (2016).

137. Chauvier, A. et al. Transcriptional pausing at the translation start site operates as a critical checkpoint for riboswitch regulation. *Nat Commun* **8**, 13892 (2017).

138. Yakhnin, A.V., Yakhnin, H. & Babitzke, P. RNA polymerase pausing regulates translation initiation by providing additional time for TRAP-RNA interaction. *Mol Cell* **24**, 547-57 (2006).
139. Mondal, S., Yakhnin, A.V., Sebastian, A., Albert, I. & Babitzke, P. NusA-dependent transcription termination prevents misregulation of global gene expression. *Nat Microbiol* **1**, 15007 (2016).
140. Yakhnin, A.V. & Babitzke, P. NusA-stimulated RNA polymerase pausing and termination participates in the Bacillus subtilis trp operon attenuation mechanism invitro. *Proc Natl Acad Sci U S A* **99**, 11067-72 (2002).
141. Gilbert, S.D. & Batey, R.T. Riboswitches: fold and function. *Chem Biol* **13**, 805-7 (2006).
142. Garst, A.D., Edwards, A.L. & Batey, R.T. Riboswitches: structures and mechanisms. *Cold Spring Harb Perspect Biol* **3**(2011).
143. Strobel, E.J., Cheng, L., Berman, K.E., Carlson, P.D. & Lucks, J.B. A ligand-gated strand displacement mechanism for ZTP riboswitch transcription control. *Nat Chem Biol* **15**, 1067-1076 (2019).
144. Wickiser, J.K., Winkler, W.C., Breaker, R.R. & Crothers, D.M. The speed of RNA transcription and metabolite binding kinetics operate an FMN riboswitch. *Mol Cell* **18**, 49-60 (2005).
145. Vassylyev, D.G., Vassylyeva, M.N., Perederina, A., Tahirov, T.H. & Artsimovitch, I. Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* **448**, 157-62 (2007).
146. Perdrizet, G.A., 2nd, Artsimovitch, I., Furman, R., Sosnick, T.R. & Pan, T. Transcriptional pausing coordinates folding of the aptamer domain and the expression platform of a riboswitch. *Proc Natl Acad Sci U S A* **109**, 3323-8 (2012).
147. Gusarov, I. & Nudler, E. Control of Intrinsic Transcription Termination by N and NusA. *Cell* **107**, 437-449 (2001).
148. Artsimovitch, I. & Landick, R. Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. *Proc Natl Acad Sci U S A* **97**, 7090-5 (2000).
149. Rinaldi, A.J., Lund, P.E., Blanco, M.R. & Walter, N.G. The Shine-Dalgarno sequence of riboswitch-regulated single mRNAs shows ligand-dependent accessibility bursts. *Nat Commun* **7**, 8976 (2016).
150. Blanco, M. & Walter, N.G. Analysis of Complex Single-Molecule FRET Time Trajectories. in *Single Molecule Tools: Fluorescence Based Approaches, Part A* 153-178 (2010).
151. Nechooshtan, G., Elgrably-Weiss, M. & Altuvia, S. Changes in transcriptional pausing modify the folding dynamics of the pH-responsive RNA element. *Nucleic Acids Res* **42**, 622-30 (2014).
152. von Hippel, P.H. Transcriptional pausing caught in the act. *Cell* **125**, 1027-8 (2006).
153. Lemay, J.F., Penedo, J.C., Tremblay, R., Lilley, D.M. & Lafontaine, D.A. Folding of the adenine riboswitch. *Chem Biol* **13**, 857-68 (2006).

154. Sung, H.L. & Nesbitt, D.J. Sequential Folding of the Nickel/Cobalt Riboswitch Is Facilitated by a Conformational Intermediate: Insights from Single-Molecule Kinetics and Thermodynamics. *J Phys Chem B* **124**, 7348-7360 (2020).
155. Breaker, R.R. Prospects for riboswitch discovery and analysis. *Mol Cell* **43**, 867-79 (2011).
156. Artsimovitch, I. & Vassylyev, D.G. Is it easy to stop RNA polymerase? *Cell Cycle* **5**, 399-404 (2006).
157. Ma, C., Yang, X. & Lewis, P.J. Bacterial Transcription as a Target for Antibacterial Drug Development. *Microbiol Mol Biol Rev* **80**, 139-60 (2016).
158. Landick, R., Wang, D. & Chan, C.L. Quantitative analysis of transcriptional pausing by Escherichia coli RNA polymerase: his leader pause site as paradigm. in *RNA Polymerase and Associated Factors, Part B* 334-353 (1996).
159. Araujo, P.R. et al. Before It Gets Started: Regulating Translation at the 5' UTR. *Comp Funct Genomics* **2012**, 475731 (2012).
160. Dvir, S. et al. Deciphering the rules by which 5'-UTR sequences affect protein expression in yeast. *Proc Natl Acad Sci U S A* **110**, E2792-801 (2013).
161. Leppek, K., Das, R. & Barna, M. Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nat Rev Mol Cell Biol* **19**, 158-174 (2018).
162. Braun, F., Durand, S. & Condon, C. Initiating ribosomes and a 5'/3'-UTR interaction control ribonuclease action to tightly couple B. subtilis hbs mRNA stability with translation. *Nucleic Acids Res* **45**, 11386-11400 (2017).
163. Kozak, M. Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene* **361**, 13-37 (2005).
164. Melnikov, S. et al. One core, two shells: bacterial and eukaryotic ribosomes. *Nat Struct Mol Biol* **19**, 560-7 (2012).
165. Na, D., Lee, S. & Lee, D. Mathematical modeling of translation initiation for the estimation of its efficiency to computationally design mRNA sequences with desired expression levels in prokaryotes. *BMC Syst Biol* **4**, 71 (2010).
166. Mustoe, A.M. et al. Pervasive Regulatory Functions of mRNA Structure Revealed by High-Resolution SHAPE Probing. *Cell* **173**, 181-195.e18 (2018).
167. Tuller, T. & Zur, H. Multiple roles of the coding sequence 5' end in gene expression regulation. *Nucleic Acids Res* **43**, 13-28 (2015).
168. de Smit, M.H. & van Duin, J. Translational Standby Sites: How Ribosomes May Deal with the Rapid Folding Kinetics of mRNA. *Journal of Molecular Biology* **331**, 737-743 (2003).
169. Mustoe, A.M., Corley, M., Laederach, A. & Weeks, K.M. Messenger RNA Structure Regulates Translation Initiation: A Mechanism Exploited from Bacteria to Humans. *Biochemistry* **57**, 3537-3539 (2018).
170. Andreeva, I., Belardinelli, R. & Rodnina, M.V. Translation initiation in bacterial polysomes through ribosome loading on a standby site on a highly translated mRNA. *Proceedings of the National Academy of Sciences* **115**, 4411-4416 (2018).

171. Shine, J. & Dalgarno, L. Terminal-sequence analysis of bacterial ribosomal RNA. Correlation between the 3'-terminal-polypyrimidine sequence of 16-S RNA and translational specificity of the ribosome. *Eur J Biochem* **57**, 221-30 (1975).

172. Ma, J., Campbell, A. & Karlin, S. Correlations between Shine-Dalgarno sequences and gene features such as predicted expression levels and operon structures. *J Bacteriol* **184**, 5733-45 (2002).

173. Simonetti, A. et al. A structural view of translation initiation in bacteria. *Cell Mol Life Sci* **66**, 423-36 (2009).

174. Del Campo, C., Bartholomaus, A., Fedyunin, I. & Ignatova, Z. Secondary Structure across the Bacterial Transcriptome Reveals Versatile Roles in mRNA Regulation and Function. *PLoS Genet* **11**, e1005613 (2015).

175. Chiaruttini, C. & Guillier, M. On the role of mRNA secondary structure in bacterial translation. *Wiley Interdiscip Rev RNA*, e1579 (2019).

176. Chen, J. et al. Coupling of mRNA Structure Rearrangement to Ribosome Movement during Bypassing of Non-coding Regions. *Cell* **163**, 1267-1280 (2015).

177. Passalacqua, K.D. et al. Structure and complexity of a bacterial transcriptome. *J Bacteriol* **191**, 3203-11 (2009).

178. Boël, G. et al. Codon influence on protein expression in E. coli correlates with mRNA levels. *Nature* **529**, 358-363 (2016).

179. Scharff, L.B., Childs, L., Walther, D. & Bock, R. Local absence of secondary structure permits translation of mRNAs that lack ribosome-binding sites. *PLoS genetics* **7**, e1002155-e1002155 (2011).

180. Winkler, W.C., Nahvi, A., Roth, A., Collins, J.A. & Breaker, R.R. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* **428**, 281-6 (2004).

181. Winkler, W.C. & Breaker, R.R. Regulation of bacterial gene expression by riboswitches. *Annu Rev Microbiol* **59**, 487-517 (2005).

182. Nudler, E. & Mironov, A.S. The riboswitch control of bacterial metabolism. *Trends Biochem Sci* **29**, 11-7 (2004).

183. Winkler, W.C., Nahvi, A., Sudarsan, N., Barrick, J.E. & Breaker, R.R. An mRNA structure that controls gene expression by binding S-adenosylmethionine. *Nat Struct Biol* **10**, 701-7 (2003).

184. Barrick, J.E. & Breaker, R.R. The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biol* **8**, R239 (2007).

185. Fuchs, R.T., Grundy, F.J. & Henkin, T.M. S-adenosylmethionine directly inhibits binding of 30S ribosomal subunits to the SMK box translational riboswitch RNA. *Proc Natl Acad Sci U S A* **104**, 4876-80 (2007).

186. Haller, A., Altman, R.B., Souliere, M.F., Blanchard, S.C. & Micura, R. Folding and ligand recognition of the TPP riboswitch aptamer at single-molecule resolution. *Proc Natl Acad Sci U S A* **110**, 4188-93 (2013).

187. Eichhorn, C.D., Kang, M. & Feigon, J. Structure and function of preQ(1) riboswitches. *Biochim Biophys Acta* **1839**, 939-950 (2014).

188. Kim, J.N. & Breaker, R.R. Purine sensing by riboswitches. *Biol Cell* **100**, 1-11 (2008).

189. Jenkins, J.L., Krucinska, J., McCarty, R.M., Bandarian, V. & Wedekind, J.E. Comparison of a preQ1 riboswitch aptamer in metabolite-bound and free states with implications for gene regulation. *J Biol Chem* **286**, 24626-37 (2011).

190. Liberman, J.A., Bogue, J.T., Jenkins, J.L., Salim, M. & Wedekind, J.E. ITC analysis of ligand binding to preQ$_1$ riboswitches. *Methods Enzymol* **549**, 435-50 (2014).

191. Liberman, J.A., Salim, M., Krucinska, J. & Wedekind, J.E. Structure of a class II preQ1 riboswitch reveals ligand recognition by a new fold. *Nat Chem Biol* **9**, 353-5 (2013).

192. Coppins, R.L., Hall, K.B. & Groisman, E.A. The intricate world of riboswitches. *Current opinion in microbiology* **10**, 176-181 (2007).

193. Ray, S., Chauvier, A. & Walter, N.G. Kinetics coming into focus: single-molecule microscopy of riboswitch dynamics. *RNA Biology* **16**, 1077-1085 (2019).

194. Lund, P.E., Chatterjee, S., Daher, M. & Walter, N.G. Protein unties the pseudoknot: S1-mediated unfolding of RNA higher order structure. *Nucleic Acids Res* (2019).

195. Culver, G.M. Meanderings of the mRNA through the ribosome. *Structure* **9**, 751-8 (2001).

196. Rio, D.C. Filter-binding assay for analysis of RNA-protein interactions. *Cold Spring Harb Protoc* **2012**, 1078-81 (2012).

197. Dorywalska, M. et al. Site-specific labeling of the ribosome for single-molecule spectroscopy. *Nucleic Acids Res* **33**, 182-9 (2005).

198. Blanchard, S.C., Kim, H.D., Gonzalez, R.L., Jr., Puglisi, J.D. & Chu, S. tRNA dynamics on the ribosome during translation. *Proc Natl Acad Sci U S A* **101**, 12893-8 (2004).

199. Francis, K. & Kohen, A. Standards for the reporting of kinetic isotope effects in enzymology. *Perspectives in Science* **1**, 110-120 (2014).

200. Li, A., Ziehr, J.L. & Johnson, K.A. A new general method for simultaneous fitting of temperature and concentration dependence of reaction rates yields kinetic and thermodynamic parameters for HIV reverse transcriptase specificity. *The Journal of biological chemistry* **292**, 6695-6702 (2017).

201. Espah Borujeni, A. & Salis, H.M. Translation Initiation is Controlled by RNA Folding Kinetics via a Ribosome Drafting Mechanism. *Journal of the American Chemical Society* **138**, 7016-7023 (2016).

202. Spitzer, J. & Poolman, B. The Role of Biomacromolecular Crowding, Ionic Strength, and Physicochemical Gradients in the Complexities of Life's Emergence. *Microbiology and Molecular Biology Reviews* **73**, 371 (2009).

203. Qu, X. et al. Single-molecule nonequilibrium periodic Mg²⁺-concentration jump experiments reveal details of the early folding pathways of a large RNA. *Proceedings of the National Academy of Sciences* **105**, 6602-6607 (2008).

204. Omotajo, D., Tate, T., Cho, H. & Choudhary, M. Distribution and diversity of ribosome binding sites in prokaryotic genomes. *BMC Genomics* **16**, 604 (2015).

205. Nakagawa, S., Niimura, Y., Miura, K.-i. & Gojobori, T. Dynamic evolution of translation initiation mechanisms in prokaryotes. *Proceedings of the National Academy of Sciences* **107**, 6382-6387 (2010).

206. Tsai, A. et al. Heterogeneous pathways and timing of factor departure during translation initiation. *Nature* **487**, 390-393 (2012).

207. Fei, J. et al. Chapter 12 - A Highly Purified, Fluorescently Labeled In Vitro Translation System for Single-Molecule Studies of Protein Synthesis. in *Methods in Enzymology*, Vol. 472 (ed. Walter, N.G.) 221-259 (Academic Press, 2010).

208. MacDougall, D.D. & Gonzalez, R.L., Jr. Translation initiation factor 3 regulates switching between different modes of ribosomal subunit joining. *J Mol Biol* **427**, 1801-18 (2015).

209. Wang, J., Caban, K. & Gonzalez, R.L., Jr. Ribosomal initiation complex-driven changes in the stability and dynamics of initiation factor 2 regulate the fidelity of translation initiation. *Journal of molecular biology* **427**, 1819-1834 (2015).

210. Tollerson, R., 2nd & Ibba, M. Translational regulation of environmental adaptation in bacteria. *J Biol Chem* **295**, 10434-10445 (2020).

211. Wang, C. et al. Structural basis of transcription-translation coupling. *Science* **369**, 1359-1365 (2020).

212. Webster, M.W. et al. Structural basis of transcription-translation coupling and collision in bacteria. *Science* **369**, 1355-1359 (2020).

213. Fletcher, S.P., Ali, I.K., Kaminski, A.N.N., Digard, P. & Jackson, R.J. The influence of viral coding sequences on pestivirus IRES activity reveals further parallels with translation initiation in prokaryotes. *RNA* **8**, 1558-1571 (2002).

214. Morris, M.J. & Basu, S. An unusually stable G-quadruplex within the 5'-UTR of the MT3 matrix metalloproteinase mRNA represses translation in eukaryotic cells. *Biochemistry* **48**, 5313-9 (2009).

215. Morris, M.J., Negishi, Y., Pazsint, C., Schonhoft, J.D. & Basu, S. An RNA G-Quadruplex Is Essential for Cap-Independent Translation Initiation in Human VEGF IRES. *Journal of the American Chemical Society* **132**, 17831-17839 (2010).

216. Bastet, L. et al. Translational control and Rho-dependent transcription termination are intimately linked in riboswitch regulation. *Nucleic Acids Research* **45**, 7474-7486 (2017).

217. Azam, M.S. & Vanderpool, C.K. Translational regulation by bacterial small RNAs via an unusual Hfq-dependent mechanism. *Nucleic Acids Res* **46**, 2585-2599 (2018).

218. Caron, M.-P. et al. Dual-acting riboswitch control of translation initiation and mRNA decay. *Proceedings of the National Academy of Sciences of the United States of America* **109**, E3444-E3453 (2012).

219. Bevilacqua, P.C. & Cech, T.R. Minor-groove recognition of double-stranded RNA by the double-stranded RNA-binding domain from the RNA-activated protein kinase PKR. *Biochemistry* **35**, 9983-94 (1996).

220. Wong, I. & Lohman, T.M. A double-filter method for nitrocellulose-filter binding: application to protein-nucleic acid interactions. *Proceedings of the National Academy of Sciences* **90**, 5428-5432 (1993).
221. Chandradoss, S.D. et al. Surface passivation for single-molecule protein studies. *J Vis Exp* (2014).
222. Verveer, P.J. & Bastiaens, P.I.H. Evaluation of global analysis algorithms for single frequency fluorescence lifetime imaging microscopy data. *Journal of Microscopy* **209**, 1-7 (2003).
223. Verveer, P.J., Squire, A. & Bastiaens, P.I.H. Global Analysis of Fluorescence Lifetime Imaging Microscopy Data. *Biophysical Journal* **78**, 2127-2137 (2000).
224. Duss, O. et al. Real-time assembly of ribonucleoprotein complexes on nascent RNA transcripts. *Nature Communications* **9**, 5087 (2018).
225. CDC. Antibiotic Resistance Threats in the United States. (ed. Prevention, C.f.D.C.a.) (2019).
226. Plackett, B. Why big pharma has abandoned antibiotics. *Nature* **586**, S50-S52 (2020).
227. McCoy, L.S., Xie, Y. & Tor, Y. Antibiotics that target protein synthesis. *Wiley Interdiscip Rev RNA* **2**, 209-32 (2011).
228. Howe, J.A. et al. Selective small-molecule inhibition of an RNA structural element. *Nature* **526**, 672-7 (2015).
229. Lunse, C.E., Schuller, A. & Mayer, G. The promise of riboswitches as potential antibacterial drug targets. *Int J Med Microbiol* **304**, 79-92 (2014).
230. Bandyopadhyay, S., Chaudhury, S., Mehta, D. & Ramesh, A. Discovery of iron-sensing bacterial riboswitches. *Nat Chem Biol* (2020).
231. Xu, J. & Cotruvo, J.A., Jr. The czcD (NiCo) Riboswitch Responds to Iron(II). *Biochemistry* **59**, 1508-1516 (2020).
232. Greenleaf, W.J., Frieda, K.L., Foster, D.A., Woodside, M.T. & Block, S.M. Direct observation of hierarchical folding in single riboswitch aptamers. *Science* **319**, 630-3 (2008).
233. Anthony, P.C., Perez, C.F., Garcia-Garcia, C. & Block, S.M. Folding energy landscape of the thiamine pyrophosphate riboswitch aptamer. *Proc Natl Acad Sci U S A* **109**, 1485-9 (2012).
234. Chandra, V., Hannan, Z., Xu, H. & Mandal, M. Single-molecule analysis reveals multi-state folding of a guanine riboswitch. *Nat Chem Biol* **13**, 194-201 (2017).
235. Duesterberg, V.K., Fischer-Hwang, I.T., Perez, C.F., Hogan, D.W. & Block, S.M. Observation of long-range tertiary interactions during ligand binding by the TPP riboswitch aptamer. *Elife* **4**(2015).
236. Lutz, B., Faber, M., Verma, A., Klumpp, S. & Schug, A. Differences between cotranscriptional and free riboswitch folding. *Nucleic Acids Res* **42**, 2687-96 (2014).

237. Chauvier, A., Nadon, J.F., Grondin, J.P., Lamontagne, A.M. & Lafontaine, D.A. Role of a hairpin-stabilized pause in the Escherichia coli thiC riboswitch function. *RNA Biol* **16**, 1066-1073 (2019).
238. Gough, N.R. Switched on by metal. *Science Signaling* **8**, ec84-ec84 (2015).
239. Kohler, R., Mooney, R.A., Mills, D.J., Landick, R. & Cramer, P. Architecture of a transcribing-translating expressome. *Science* **356**, 194-197 (2017).
240. Duss, O. et al. Real-time assembly of ribonucleoprotein complexes on nascent RNA transcripts. *Nat Commun* **9**, 5087 (2018).
241. Warnasooriya, C. et al. Observation of preQ1-II riboswitch dynamics using single-molecule FRET. *RNA Biol* **16**, 1086-1092 (2019).
242. Polaski, J.T., Holmstrom, E.D., Nesbitt, D.J. & Batey, R.T. Mechanistic Insights into Cofactor-Dependent Coupling of RNA Folding and mRNA Transcription/Translation by a Cobalamin Riboswitch. *Cell Rep* **15**, 1100-1110 (2016).
243. Rinaldi, A.J., Suddala, K.C. & Walter, N.G. Native purification and labeling of RNA for single molecule fluorescence studies. *Methods Mol Biol* **1240**, 63-95 (2015).
244. Righetti, F. & Narberhaus, F. How to find RNA thermometers. *Front Cell Infect Microbiol* **4**, 132 (2014).
245. Neupert, J., Karcher, D. & Bock, R. Design of simple synthetic RNA thermometers for temperature-controlled gene expression in Escherichia coli. *Nucleic Acids Res* **36**, e124 (2008).
246. Abduljalil, J.M. Bacterial riboswitches and RNA thermometers: Nature and contributions to pathogenesis. *Noncoding RNA Res* **3**, 54-63 (2018).